University of Alberta

Analysis of the *bldA* tRNA Gene of *Streptomyces clavuligerus*: Implications for Mistranslation of TTA Codons

by

Nicole Katherine Trepanier ©

A thesis submitted to the Faculty of Graduate Studies and Research in partial fulfillment of the requirements for the degree of Doctor of Philosophy

in

Molecular Biology and Genetics

Department of Biological Sciences

Edmonton, Alberta

Fall, 1999

National Library of Canada
Acquisitions and Bibliographic Services
395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque nationale du Canada
Acquisitions et services bibliographiques
395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence

Our file Notre référence

The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

L'auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author's permission.

L'auteur conserve la propriété du droit d'auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-46935-2

Canada

University of Alberta

Library Release Form

Name of Author: Nicole Katherine Trepanier

Title of Thesis: Analysis of the *bldA* tRNA Gene of *Streptomyces clavuligerus*: Implications for Mistranslation of TTA Codons

Degree: Doctor of Philosophy

Year this Degree Granted: 1999

Permission is hereby granted to the UNIVERSITY OF ALBERTA LIBRARY to reproduce single copies of this thesis and to lend or sell such copies for private, scholarly, or scientific research purposes only.

The author reserves the other publication rights, and neither the thesis nor extensive extracts from it may be printed or otherwise reproduced without the author's written permission.

N Trepanier

(Student's signature)

Student's permanent address:

R. R. #2

Grande Prairie, Alberta

T8V 2Z9

Date: Aug 6, 1999

University of Alberta

Faculty of Graduate Studies and Research

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research for acceptance, a thesis entitled Analysis of the *bldA* tRNA Gene of *Streptomyces clavuligerus*: Implications for Mistranslation of TTA Codons submitted by Nicole Katherine Trepanier in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Molecular Biology and Genetics.

Brenda Leskiw
Dr. Brenda K. Leskiw (Supervisor)

Kenneth L. Roy
Dr. Kenneth L. Roy

Laura Frost
Dr. Laura S. Frost

SE Jensen
Dr. Susan E. Jensen

Mark Peppler
Dr. Mark Peppler

H.E.M. for
Dr. Wendy C. Champness

Date: August 9, 1999

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Dr. Brenda K. Leskiw, for her unending patience, support and guidance in all aspects of this thesis work. In particular I would like to thank her for the encouragement she has given me to express new ideas and to continue to look for answers.

I would also like to thank my supervisory committee, Dr. Kenneth L. Roy and Dr. Laura Frost. Their helpful and timely suggestions as well as the time and effort they have devoted to this project are greatly appreciated. I would also like to acknowledge both past and present members of Dr. Leskiw's laboratory, especially Marie Elliot, Leigh Ann Geibelhaus, and Janice Strap who have helped with various aspects of my project and made the lab an enjoyable place to work. The project work of several undergraduate students, Hadley Marcus, Paul Leedell, and Grace Nguyen contributed to this thesis and their individual contributions are appreciated. In addition, I would like to thank Dr. Susan E. Jensen and all members of her laboratory for their generosity in sharing reagents, equipment and protocols for work with *Streptomyces clavuligerus*. In particular, I wish to acknowledge the technical assistance provided by Dr. Ashish Paradkar as well as the help of Annie Wong, Cecilia Anders, and Wendy Thai. Thanks also go to the Prep Room staff; Tom Hantos, Elsa Bruno, and Anna Szenthe as well as to the members of Dr. Frost's lab who also shared reagents, equipment and protocols with me.

Words cannot adequately express my appreciation for the help and support provided by my friends and family. Without the friendship of Jeff Hoyem, Elsa Bruno, and Marie Elliot, I would not have been able to get through some of the difficult times. I would like to also thank Jeff Hoyem for his patience and assistance with my many computer problems and his uncanny ability to find and/or fix things. My parents and other family members have given me unwavering support and encouragement and believed in me even when I doubted myself. My parents in particular have shared this long and arduous journey with me and without them I would never have completed this thesis.

ABSTRACT

The *bldA* gene encodes the $\text{tRNA}^{\text{Leu}}_{\text{UAA}}$ that translates the rare UUA leucine codon in the high G +C (~70%) *Streptomyces* mRNA. These rare TTA codons are confined to *Streptomyces* genes that function relatively late in growth, after the switch from primary metabolism and vegetative growth to secondary metabolism and differentiation. *bldA* mutant strains of *S. coelicolor, S. lividans,* and *S. griseus* are defective in sporulation and antibiotic production, as key genes required for these processes contain TTA codons and are not efficiently translated in the absence of the *bldA* $\text{tRNA}^{\text{Leu}}_{\text{UAA}}$.

In an effort to gain evidence that sporulation and antibiotic production are similarly regulated by *bldA* in other (perhaps all) *Streptomyces* species, the *bldA* gene was cloned and sequenced in *Streptomyces clavuligerus,* and a *bldA* mutant strain was generated by gene replacement. The phenotype of the mutant strain was sporulation defective as expected, however, the *S. clavuligerus bldA* mutant was not defective in antibiotic production, as cephamycin C, and clavulanic acid were produced in the mutant strain. Subsequent analysis indicated that antibiotic production in the absence of the *bldA* $\text{tRNA}^{\text{Leu}}_{\text{UAA}}$ was due to the efficient mistranslation of the single TTA codon in the cephamycin C and clavulanic acid pathway-specific activator gene, *ccaR*, possibly by wobble base pairing of the $\text{tRNA}^{\text{Leu}}_{\text{CAA}}$ with the UUA codon.

Since a TTA-containing reporter gene was not expressed in the *S. clavuligerus bldA* mutant it was evident that not all UUA codons are mistranslated efficiently in the *S. clavuligerus bldA* mutant strain, and that there must be something unique about the context of the TTA codon in the *ccaR* gene that makes its expression essentially independent of the *bldA* tRNA. When the *ccaR* gene was compared to several TTA-

containing genes that were shown to be dependent on *bldA* for their expression, it was found that *ccaR* contains a guanine nucleotide 3' of the TTA codon whereas most TTA codons in other genes contain a cytidine in this position. This observation lead to the hypothesis that +1 frameshifts are introduced at TTA C and TTA T sequences by the $\mathrm{tRNA}^{\mathrm{Tyr}}_{\mathrm{GUA}}$ in *bldA* mutants while TTA G and TTA A sequences are mistranslated in-frame, as +1 shifts are prevented by the TAG or TAA stop codon in the +1 frame.

TABLE OF CONTENTS

Page

Abstract

Acknowledgements

List of Tables

List of Figures

List of Abbreviations

I INTRODUCTION 1

II. MATERIALS AND METHODS 59

II.1 MATERIALS 59

II.2 BACTERIAL STRAINS, PLASMIDS, PHAGES AND CULTURE CONDITIONS 63

- II.2.1 Bacterial strains, plasmids and phages 63
- II.2.2 Maintenance of *Streptomyces* strains 63
- II.2.3 Maintenance of *E. coli* and *S. aureus* strains 68
- II.2.4 Culture conditions for isolation of DNA 68
- II.2.5 Culture conditions for preparation of cell free extracts 70
- II.2.6 Culture conditions for preparation of protoplasts 70
- II.2.7 Culture condition for isolation of RNA 71

II.3 ISOLATION OF CHROMOSOMAL DNA, PLASMID DNA AND RNA 72

- II.3.1 Isolation of chromosomal, plasmid, and single stranded DNA 72
- II.3.2 Isolation of RNA 73

II.4 INTRODUCING DNA BY TRANSFORMATION 74

- II.4.1 Transformation of *S. coelicolor* and *S. lividans* 74
- II.4.2 Transformation of *S. clavuligerus* 75
- II.4.3 Transformation of *E. coli* 77

II.4.4 Digestion, cloning and subcloning of DNA 78

II.5 TECHNIQUES USED FOR DNA ANALYSIS 79

II.5.1 Restriction fragment analysis by agarose gel electrophoresis 79

II.5.2 Restriction fragment analysis by polyacrylamide gel electrophoresis 79

II.5.3 DNA sequence analysis 79

II.5.4 Hybridization analysis 80

II.5.4.1 Preparation of membranes for Colony Hybridization 80

II.5.4.2 Preparation of Dot Blots 81

II.5.4.3 Preparation of membranes for Southern Hybridization 81

II.5.4.4 Generation of ^{32}P-labelled probes 82

II.5.4.5 Filter hybridization 83

II.6 RECOVERY OF DNA FROM AGAROSE AND POLYACRYLAMIDE GELS 85

II.6.1 Recovery of DNA from agarose gels 85

II.6.2 Recovery of DNA from polyacrylamide gels 85

II.7 POLYMERASE CHAIN REACTION (PCR) AND INVERSE PCR 86

II.7.1 Polymerase Chain Reaction 86

II.7.2 Inverse Polymerase Chain Reaction 86

II.8 TECHNIQUES FOR RNA ANALYSIS 87

II.8.1 Denaturing polyacrylamide gel electrophoresis 87

II.8.2 RNA agarose gel electrophoresis 87

II.8.3 RNA electroblotting onto nylon membranes from polyacrylamide gels 88

II.8.4 Northern transfer of RNA from agarose gels 88

II.8.5 Generation of ^{32}P-labelled probes 89

II.8.6 Northern analysis 89

II.8.7 *S1* nuclease protection assays 90

II.8.8 Primer Extension 91

II.9 WESTERN ANALYSIS 92

II.10 REPORTER GENE ANALYSIS 93

II.10.1 Promoter probe analysis using the *xylE* reporter gene 93

II.10.2 Catechol dioxygenase assays of *S. lividans* TK24 and *S. clavuligerus* NRRL3585 transformants containing recombinant pIJ4083 plasmids 94

II.10.3 Green Fluorescent Protein reporter gene analysis 95

II.11 ANTIBIOTIC BIOASSAYS 96

III. RESULTS 98

III.1 INITIAL SCREENING OF β-LACTAM PRODUCING *Streptomyces* STRAINS 98

III.2 CLONING OF THE *S. clavuligerus bldA* GENE 101

III.3 SUBCLONING OF THE *S. clavuligerus bldA* GENE 107

III.4 SEQUENCE ANALYSIS OF THE 1.5 kb *Sal*I FRAGMENT 107

III.5 GENERATION OF A *bldA* NULL MUTANT STRAIN 135

III.6 COMPLEMENTATION OF THE *bldA* MUTANT STRAIN 150

III.7 IDENTIFICATION OF THE *S. coelicolor leuU* tRNA GENE 166

III.8 NORTHERN ANALYSIS OF THE *S. coelicolor bldA* AND *leuU* tRNA TRANSCRIPTS 178

III.9 ANTIBIOTIC PRODUCTION AND *bldA* tRNA EXPRESSION IN *S. clavuligerus* 182

III.10 REPORTER GENE ANALYSIS 198

III.11 S1 NUCLEASE PROTECTION ANALYSIS 206

III.12 PRIMER EXTENSION ANALYSIS 214

III.13 PROMOTER PROBE ANALYSIS 218

IV. DISCUSSION 227

V. BIBLIOGRAPHY 283

VI. APPENDICES 308

LIST OF TABLES

Table		Page
I.1	*Streptomyces* genes containing TTA codons (*bldA* targets)	34
II.1.1	Oligonucleotide primer sequences	61
II.2.1.1	*Streptomyces* Strains	64
II.2.1.2	*Escherichia coli* and *Staphylococcus aureus* Strains	65
II.2.1.3	*E. coli* and *Streptomyces* plasmids	66
III.4.1	Comparison of *Streptomyces* tRNA gene sequences to identify invariant and semi-invariant bases	116
III.4.2	Invariant and semi-invariant bases in *Streptomyces* tRNA genes	118
III.6.1	Expeceted size of hybridizing *S. clavuligerus* chromosomal DNA fragments to the *bldA* and *tsr* probe	167
III.9.1	Antibiotic bioassay of surface-grown cultures of *S. clavuligerus* NRRL 3585 (wild-type) and *S. clavuligerus* Δ*bldA*4-1	183
III.9.2	Antibiotic bioassay of broth cultures of *S. clavuligerus* NRRL 3585 (wild-type) and *S. clavuligerus* Δ*bldA*4-1	189
III.13.1	Catechol dioxygenase (C23O) activity in cell free extracts of *S. clavuligerus* and *S. lividans* transformed with pIJ4083 recombinant plasmids for identification of the *bldA* promoter	223
IV.1	Wobble base pairing between the first position of the anticodon and the third codon position	244
IV.2	Comparison of *ccaR (bldA* independent target) to *bldA* dependent targets	257
IV.3	Classification of *Streptomyces* genes containing TTA codons	267
IV.4	Comparison of codon context in TTA-containing reporter genes	271
IV.5	Distribution of nucleotides following leucine codons in *Streptomyces* genes containing TTA codons	275

LIST OF FIGURES

Figure		Page
I.1	Lifecycle of *Streptomyces coelicolor*	4
I.2	Structure of β-lactam antibiotics	42
I.3	*Streptomyces clavuligerus* cephamycin C biosynthetic gene cluster	46
I.4	Biosynthetic pathway for cephamycin C	48
I.5	Structure of clavulanic acid clavam compounds	51
I.6	*Streptomyces clavuligerus* clavulanic acid biosynthetic gene cluster	54
I.7	Biosynthetic pathway for clavulanic acid	56
III.1.1	Hybridization of the *S. coelicolor bldA* gene to *S. clavuligerus*, *S. jumonjinensis*, and *S. lipmanii* DNA	100
III.2.1	Hybridization of the *S. coelicolor bldA* gene to pLAFR-11C1 cosmid DNA digested with a variety of restriction endonucleases	104
III.2.2	Hybridization of the *S. clavuligerus* 1.5 kb *Sal*I restriction fragment to *S. clavuligerus* chromosomal DNA	106
III.3.1	Restriction map of the 1.5 kb *S. clavuligerus Sal*I insert in p9S+	109
III.4.1	Sequence of the *S. clavuligerus* 1.5 kb *Sal*I fragment containing the *bldA* tRNA gene	111
III.4.2	The deduced secondary structure of the *S. clavuligerus* and *S. coelicolor bldA* tRNAs	115
III.4.3	Frame analysis of the 1.5 kb *Sal*I fragment containing the *bldA* gene	122
III.4.4	Comparison of the nucleotide and amino acid sequence of *S. clavuligerus* ORF1 and *S. coelicolor* ORFJI1	125
III.4.5	Corrected sequence of *S. coelicolor* ORFJI1	127
III.4.6	Additional sequence of *S. clavuligerus* ORF1	130

III.4.7 Alignment of *S. clavuligerus* ORF1 and *S. coelicolor* ORFJI1 amino acid sequence 132

III.4.8 Comparison of the amino acid sequence of *S. clavuligerus* ORF2 and *S. griseus* ORF 134

III.4.9 13-20 base direct repeats in 80% G+C region from base 772-1021 of the 1.5 kb *SalI bldA*-containing cloned DNA 137

III.5.1 Subcloning strategy for the generation of the *bldA* gene replacement vector 139

III.5.2 Flow diagram outlining the procedure for generation of the *S. clavuligerus ΔbldA* mutant strain 144

III.5.3 Hybridization analysis of *S. clavuligerus bldA* mutant strains 147

III.5.4 Hybridization analysis of *S. clavuligerus ΔbldA* mutants regenerated from spores 149

III.6.1 Subcloning of the *S. clavuligerus bldA* gene into the *Streptomyces* vector pIJ486 for complementation of the *bldA* mutant phenotype 152

III.6.2 Complementation of the *S. lividans* J1725 *bldA* mutant146 phenotype by the *S. clavuligerus bldA* gene on the high copy number plasmid pIJ486 155

III.6.3 Complementation of the *S. clavuligerus bldA* mutant phenotype by the *S. clavuligerus bldA* gene in single copy and on the high copy number plasmid pIJ486 157

III.6.4 Hybridization of the *S. clavuligerus bldA* gene to DNA isolated from complemented strains of *S. clavuligerus* 159

III.6.5 Hybridization of the thiostrepton resistance gene to DNA isolated from complemented strains of *S. clavuligerus* 161

III.6.6 Subcloning of the *S. clavuligerus bldA* gene into the integrating *Streptomyces* vector pSET152 for complementation of the *bldA* mutant phenotype 165

III.7.1 Hybridization of chromosomal DNA from *S. coelicolor* and *S. clavuligerus* with the ^{32}P-labelled primer BKL42 to detect leucyl-tRNAs that contain the anticodon GAG 171

III.7.2 Nucleotide sequence of the *S. coelicolor leuUα* tRNA gene and flanking sequence 174

III.7.3 Deduced secondary structure of the *S. coelicolor leuUα* tRNA 177

III.8.1 Northern blot analysis of the *S. coelicolor bldA* tRNA and *leuU* tRNA in RNA samples isolated from surface-grown cultures of *S. coelicolor* J1501. 180

III.9.1 Northern blot analysis of the *S. clavuligerus bldA* tRNA in RNA samples isolated from surface-grown cultures of *S. clavuligerus ΔbldA* and the isogenic parental strain *S. clavuligerus* NRRL 3585 (wild-type) 186

III.9.2 Northern blot analysis of the *S. clavuligerus bldA* in RNA samples isolated from liquid cultures of *S. clavuligerus ΔbldA* and the isogenic parental strain *S. clavuligerus* NRRL 3585 (wild-type) 191

III.9.3 Schematic diagram of the cephamycin C and clavulanic acid gene super cluster and the gene fragments and oligonucleotide (probe A) that were used to probe northern membranes 194

III.9.4 Northern blot analysis of the *S. clavuligerus ccaR, pcbC,* and *cas2* transcripts in RNA samples isolated from liquid cultures of *S. clavuligerus ΔbldA* and the isogenic parental strain *S. clavuligerus* NRRL 3585 (wild-type) 196

III.9.5 Western analysis of the *S. clavuligerus* CcaR protein in cell free extracts from *S. clavuligerus* NRRL 3585 (wild-type), *S. clavuligerus ΔbldA*4-1, and *S. clavuligerus ccaR::apr* isolated after 24 and 48 hours of growth 200

III.10.1 Sequence of the *mut3gfp* reporter gene containing three TTA codon 203

III.10.2 Expression of the GFP reporter gene in *S. clavuligerus* NRRL 3585 (wild-type) and *S. clavuligerus ΔbldA*4-1 205

III.11.1 Alignment of the *S. coelicolor bldA* and *S. clavuligerus bldA* tRNA genes and promoter regions 208

III.11.2 Probes for S1 nuclease protection analysis 210

III.11.3 S1 nuclease protection analysis of *S. clavuligerus bldA* transcripts 213

III.12.1 Primers for primer extension analysis 216

III.13.1 Promoter probe plasmids for identification of the *bldA* promoter 220

III.13.2 Visual assessment of catechol dioxygenase activity in patches of *S. clavuligerus* transformed with the recombinant promoter probe plasmids 226

IV.1 Cloverleaf structure of *Streptomyces* tRNAs showing invariant and semi-invariant bases 230

IV.2 Schematic diagram of the *S. clavuligerus bldA* region with the putative promoter consensus sequences indicated 233

IV.3 Production of undecylprodigiosin (red pigment) by *S. coelicolor* J1681 (A) and *S.lividans* J1725 (B) *bldA* mutants 264

IV.4 Translation of UUA codons in *Streptomyces* according to the "mistranslation theory". (A) *bldA* mutant (B) wild-type 278

LIST OF ABBREVIATIONS

CHEMICALS, REAGENTS, MEDIA, AND BUFFERS

DEPC	Diethyl pryrocarbonate
DTT	Dithiothreitol
EDTA	Ethylenediaminetetraacetic acid
IPTG	Isopropyl β-D-thiogalactopyranoside
ISP#3	International *Streptomyces* Project Medium #3
LB	Luria Bertani Medium
Modified R5B	Modified Bailey's Sucrose Medium
MYM	Malt Extract-Yeast Extract-Maltose Medium
P Buffer	Protoplast Buffer
PEG	Polyethylene glycol
R2YE	Sucrose Yeast Extract Medium
SDS	Sodium dodecyl sulfate
SSC	Standard Saline-Citrate
TAE	Tris-acetate-EDTA buffer
TBE	Tris-borate-EDTA buffer
TBS	Tris Buffered Saline
TDE	Tris-DDT-EDTA Buffer
TE	Tris-EDTA Buffer
TEMED	Tetramethyl ethylene diamine
TES	N-tris [Hydroxymethyl] methyl-2-aminoethane sulfonic acid
TOA	Tomato Oatmeal Agar
TSB	Trypticase Soy Broth
TTE	Tris Taurine EDTA Buffer
X-gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
YEME	Yeast Extract-Malt Extract Medium
YT	Yeast Extract Tryptone Medium

NUCLEOTIDES AND NUCLEOSIDES

NTP	Nucleoside triphosphate
dNTP	Deoxynucleoside triphosphate
ddNTP	Dideoxynucleside triphosphate
A	Adenine
C	Cytidine

G	Guanosine
N	Nucleotide or Nucleoside
R	Purine (G and/or A)
T	Thymidine
U	Uridine
Y	Pyrimidine (C, T and/or U)

GENES, PROTEINS AND RNA

amp	Ampicillin resistance gene
AMV	Avian Myeloblastosis Virus
apr	Apramycin resistance gene
apr^R	Apramycin resistant phenotype (same nomenclature used for resistance to other antibiotics)
apr^S	Apramcyin sensitive phenotype (same nomenclature used for sensitivity to other antibiotics)
cas2	Clavaminate Synthase 2 gene
ccaR	Cephamycin and Clavulanic Acid Regulatory gene (protein = CcaR)
ccar::apr	Apramycin resistance gene inserted into *ccaR* gene
C23O	Catechol-2,3-dioxygenase
D Stem-loop	tRNA arm that contains invariant Dihydrouridine base
ermE	Erythromycin resistance gene (or promoter)
gfp	Green Fluorescent Protein gene (protein = GFP)
glyT	Gene designation for first tRNA^{Gly} gene identified in *Streptomyces,* subsequent tRNA^{Gly} species designated *glyU,-V*, etc., homologues of *glyT* in same species designated *glyT*β, γ, etc., with first *glyT* identified designated *glyT*α. (Same nomenclature used for other tRNA species)
LeuRS	Leucyl-tRNA synthetase
pcbC	Gene encoding isopenicillin synthase
ORF	Open Reading Frame
rDNA	portion of rRNA gene corresponding to the structural rRNA
tDNA	portion of tRNA gene corresponding to the structural tRNA
tDNA$^{Leu}_{UAA}$	tDNA for leucyl-tRNA with UAA anticodon
T Stem-loop	(also TΨC) tRNA arm that contains invariant TΨC sequence
tRNA$^{Leu}_{UAA}$	tRNA with UAA anticodon that is charged with the amino acid leucine (Same designation is used for all charged tRNAs)
tsr	Thiostrepton resistance gene
xylE	Catechol dioxygenase gene

MISCELLELANEOUS MOLECULAR BIOLOGICAL TERMS

A_{260}	Absorbance at 260 nm
α	alpha (anti when referring to antibodies, eg., α-CcaR)
bp	Base pair
CFU	Colony Forming Units
Ci	Curie (radioactivity unit)
CPM	Counts per Minute (radioactive decay)
Δ	Deletion
Ig	Immunoglobulin
kb	kilobase
λ	Lambda Bacteriophage
MCS	Multiple Cloning Site
MW	Molecular Weight
OD_{600}	Optical Density (turbidity) at 600 nm
T_d	Dissociation temperature of short DNA duplexes (shorter than 50 nucleotides) –typically refers to dissociation temperature of oligonucleotide probes (temperature at which 50% of the duplexes dissociate)
T_m	Melting temperature of DNA hybrids – typically refers to melting temperature of DNA probe 50 nucleotides or longer (temperature at which half of the hybrids are dissociated)
U	Unit (measure of enzyme activity)

INTRODUCTION

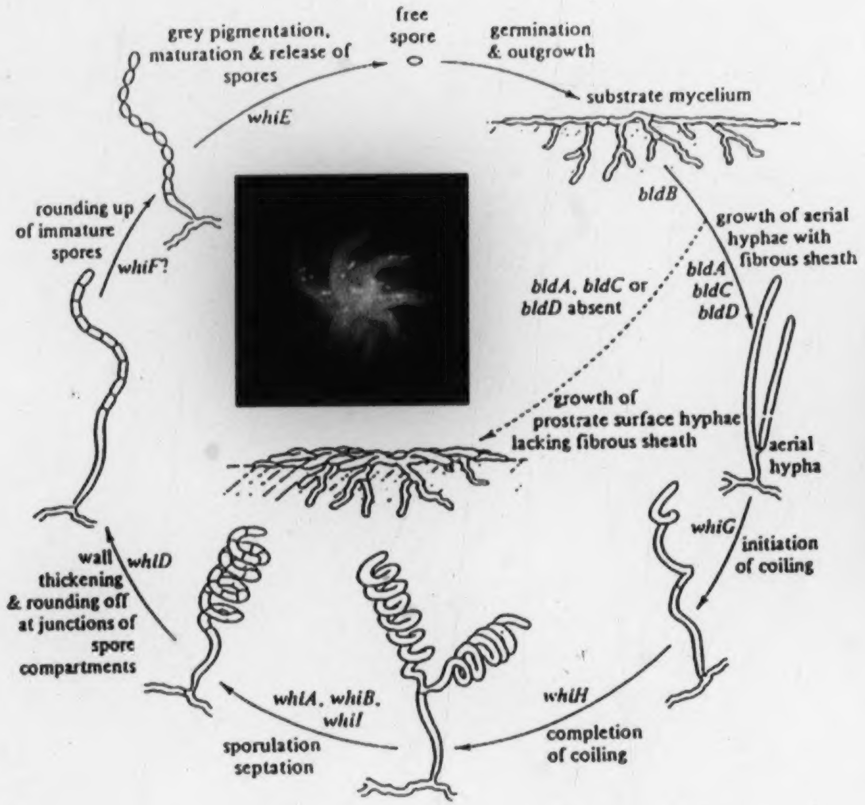
Streptomycetes are a family of Gram positive aerobic soil-dwelling bacteria that undergo a complex cycle of morphological and physiological development. They are well adapted for their ecological niche in a number of ways: they produce a variety of extracellular enzymes which enable them to degrade both plant and animal material, including polysaccharides, proteins, as well as aromatic compounds (Korn-Wendisch and Kutzner, 1992); they are able to utilize inorganic nitrogen sources; they produce spores which enable them to withstand unfavourable environmental conditions, and they produce an astounding array of secondary metabolites such as antibiotics which enable them to compete with other bacteria for nutrients. These adaptations have generated significant interest in the scientific community as these organisms are useful hosts for the study of extracellular enzyme production for the purpose of bioremediation, and medically and industrially useful organisms for the production of biologically active compounds. *Streptomyces* are also useful as a relatively simple model for the study of the process of differentiation, as they are one of the few groups of bacteria that display a complex colonial mode of growth with cellular differentiation.

Growth of *Streptomyces* on solid media begins with the germination and outgrowth of a single spore. During vegetative growth a mat of branching multinucleated substrate mycelia is formed. In response to as yet unidentified signals, the sporulation process is initiated. This process is characterized by three features: electron transparent granules, which may contain glycogen, are produced in the cytoplasm of the substrate mycelium; aerial hyphae begin to grow, and antibiotic production is initiated (Chater, 1989; Chater, 1993). While the aerial hyphae are developing, the substrate mycelia are cannibalized to provide nutrients for the developing hyphae. It is believed that the breakdown of the electron transparent granules, in addition to providing nutrients allowing for extension of the aerial hyphae, also maintains turgor to keep the hyphae extended into the air (Chater, 1989; Plaskitt and Chater, 1995). The aerial hyphae coil, septate, and round off to form chains of pigmented, uninucleate, rounded spores.

The sequence of events during the formation of aerial hyphae and maturation into spore chains, as well as some of the specific signals regulating differentiation, have been elucidated with the isolation and characterization of mutants defective in the sporulation process. The first *Streptomyces* developmental mutants identified were in the model streptomycete, *Streptomyces coelicolor* (Hopwood *et al.*, 1970). These were designated "white" mutants, as the mutant colonies failed to form chains of grey-pigmented spores and retained the white appearance imparted by the aerial mycelia. Six subclasses of *whi* (white) mutants have been identified and mapped to nine loci (*whi A-E,-G-J)* (Chater, 1972) (Chater, personal communication). Ultrastructural studies have shown that each of these mutants appears to be blocked at a specific point in the maturation of the aerial hyphae into pigmented spore chains (Hopwood *et al.*, 1970; McVittie, 1974), and the mutant phenotypes have been useful in defining the sequence of events from the appearance of aerial hyphae to the generation of mature, pigmented, uninucleate, rounded spores (Figure I.1). Based on the phenotype of the various *whi* mutants, as well as phenotype of double *whi* mutants (Chater, 1975), the temporal order of function of the *whi* genes is thought to occur in the order *whiG, whiJ* < *whiA, whiB* < *whiH* < *whiI* < *whiD, whiE. whiG* encodes a σ factor that is closely related to the motility σ factor of *Bacillus subtilis* (Chater *et al.*, 1989) and *whiG* transcripts are present in all of the other

Figure I.1

Lifecycle of *Streptomyces coelicolor*. The *bld* genes are implicated in the production of aerial hyphae, and the *whi* genes are implicated in the maturation of aerial hyphae into chains of pigmented spores. The temporal order of the various *bld* and *whi* genes indicated is based on the phenotype of the respective mutant strains. **A)** *bldD* mutant displaying the shiny appearance characteristic of *bld* mutants. **B)** A wild-type strain showing the white appearance characteristic of immature aerial hyphae as well as some red pigmentation due to the production of the cell-associated antibiotic undecylprodigiosin. Reprinted from Chater and Merrick (1979) with modifications.


free spore
germination & outgrowth
substrate mycelium
bldB
growth of aerial hyphae with fibrous sheath
bldA
bldC
bldD
bldA, *bldC* or *bldD* absent
growth of prostrate surface hyphae lacking fibrous sheath
aerial hypha
whiG
initiation of coiling
whiH
completion of coiling
whiA, *whiB*, *whiI*
sporulation septation
whiD
wall thickening & rounding off at junctions of spore compartments
whiF?
rounding up of immature spores
whiE
grey pigmentation, maturation & release of spores

whi mutant strains confirming its function early in sporulation. The *whiJ* mutation is complemented by a DNA fragment containing four genes, two of which encode proteins related to the products of *abaA (a*ntibiotic *b*iosynthesis *a*ctivator) locus (Fernández-Moreno *et al.*, 1992), although it has not yet been determined which of the genes in the complementing fragment is required for WhiJ function (K. F. Chater, personal communication). *whiA* encodes a protein that shows homology to a mycobacterial putative transcription factor (Ryding and Chater, personal communication). *whiB* encodes a small, highly charged protein which may function as a transcription factor (Davis and Chater, 1992). The *whiH* gene, which contains a *whiG*-dependent promoter, encodes a protein that is related to the GntR family of repressors and transcriptional activators, which often regulate aspects of carbon metabolism (Ryding *et al.*, 1998). *whiE* maps to an operon of at least eight genes that are involved in the biosynthesis of the grey spore color (Davis and Chater, 1990; Kelemen *et al.*, 1998) and the transcription of the *whiE* locus is regulated by two promoters that depend on the products of the *whiA, whiB, whiG, whiI,* and to a lesser extent *whiH* and *whiJ* genes for transcription. The *whiI* locus has recently been identified as two genes that belong to the family of two-component regulators consisting of a sensor-kinase and response regulator (K. Chater, John Innes Centre http://www.jic.bbsrc.ac.uk/).

Another mutation that imparts a *whi* phenotype is *sigF* from *Streptomyces aureofaciens* and *S. coelicolor.* The *sigF* gene encodes a σ factor that shows homology to the σ^B factor of *Bacillus subtilis* which is involved in stationary-phase gene expression (Potuckova *et al.*, 1995) and the absence of expression of *sigF* in most of the *whi* mutant strains (*whiA,- B, -G* and *-I)* and reduced expression in *whiH* and *whiJ* strains indicates that this gene functions late in sporulation (Kelemen *et al.*, 1996).

The *esp* (ectopic sporulation) locus of *S. coelicolor* was discovered in several strains containing deletions of the glucose kinase gene (*glkA)* (Kelemen *et al.*, 1995). These strains formed spores ectopically (on vegetative hyphae), a phenotype similar to that observed when the *whiG* sigma factor is overexpressed (Chater *et al.*, 1989). Further analysis of these mutant strains indicated that an additional gene that was deleted along with the *glkA* gene was responsible for the *esp* phenotype. An *espwhiG* double mutant displayed a *whiG* phenotype indicating that the *esp* gene product likely interacts with WhiG and regulates its temporal and spatial expression. At present, the gene has not yet been identified, although it is speculated that it could encode an anti-sigma factor that interacts with WhiG. Alternately, it may encode a repressor-type protein that acts on the *whiG* promoter, or its effect on *whiG* may be indirect.

With the exception of the *whiE* locus, all of the characterized *whi* genes have a regulatory function, indicating that sporulation in *Streptomyces* is regulated by a cascade. Two of the genes identified, *sigF* and *whiG,* encode sigma factors, suggesting that a sigma factor cascade may also be involved in sporulation in *Streptomyces* (Potuckova *et al.*, 1995), similar to the sigma factor cascade involved in sporulation in *Bacillus subtilis* (Losick and Pero, 1981).

While the sporulation process has only been intensively studied in *Streptomyces coelicolor,* it is likely that sporulation in other *Streptomyces* species will be regulated similarly. Hybridization of the *whiG* and *whiB* genes to DNA from other *Streptomyces* species and representative genera of other actinomycetes indicate that these two genes are conserved among *Streptomyces* species as well as some of the related genera (Chater, 1993). This indicates that these two genes may be required for sporulation in many of the actinomycetes, although the function of *whiB* may be more generalized since it hybridizes to DNA from actinomycetes that do not sporulate. *whiG* homologues have been identified in *Streptomyces aureofaciens* (Chater, 1993) and in *Streptoverticillum griseocarneum* (Soliveri *et al*., 1993) a member of a genus closely related to *Streptomyces*, and a *whiB* homologue has been identified in *Streptoverticillum griseocarneum* (Soliveri *et al*., 1993). Additional testing is needed to determine if other *whi* genes are similarly conserved among the streptomycetes.

The second class of developmental mutants characterized in *S. coelicolor* were found to be defective in the production of aerial hyphae and were termed "bald" as the mutant colonies maintained the shiny and bald appearance characteristic of vegetative growth (Merrick, 1976). Ultrastructural studies, however, have revealed that some of the bald *(bld)* mutants do in fact produce aerial hyphae but these hyphae lack the necessary turgor to remain erect, and these prostrate aerial hyphae fail to develop into chains of spores. Numerous *bld* mutants have been identified: *bldA-I,-K* (Chater, 1993; Puglia and Cappelletti, 1984), and several of the *bld* genes have been at least partially characterized. *bldA* has been found to encode a leucyl-tRNA that translates the rare UUA codon in *Streptomyces* G+C rich mRNA (70%) (Lawlor *et al*., 1987). *bldB* (Pope *et al*., 1998) and *bldD* (Elliot *et al*., 1998) both encode small proteins with helix-turn-helix motifs and may function as DNA-binding transcription factors. Analysis of *bldD* transcripts and *bldB* promoter dependent expression suggest that both of these genes are auto-regulated, as transcription from both promoters is increased in the respective mutant strains.

Preliminary work has indicated that *bldG* may encode an anti-anti-sigma factor (Leskiw, personal communication), which predicts the involvement of an additional anti-sigma and sigma factor in the regulation of differentiation. *bldK* has been identified as a gene cluster containing five open reading frames (*bldKA-E*) that encode homologues of the polypeptide components of the ATP-binding cassette (ABC) membrane spanning transporters (Nodwell *et al.*, 1996), and likely functions as an oligopeptide import system. Although the *bldI* gene has not yet been characterized, the level of *bldA* transcripts in *bldI* mutants is significantly reduced indicating that *bldI* may regulate *bldA* expression, either directly or indirectly (Leskiw and Mah, 1995).

The defect in aerial mycelium formation for most of the *bld* mutants is conditional, when grown on glucose containing media the mutant phenotype is observed but when grown on alternate carbon sources, such as mannitol, sporulation is restored (Champness, 1988; Merrick, 1976). Exceptions are *bldB* and *bldI* which sporulate on rich media after prolonged incubation, and *bldC* and *bldF* which are unconditionally blocked in aerial mycelia formation.

Unlike the *whi* mutants, which are only defective in sporulation, *bld* mutants are pleiotropic, as most of them are also defective in antibiotic production. *bldA,-B,-D,-G, -H,* and *-I* are defective in production of all four of the antibiotics produced by *S. coelicolor* (actinorhodin, undecylprodigiosin, methylenomycin, and calcium dependent antibiotic), although *bldH* mutants are conditionally defective for all four antibiotics, *bldA* mutants are conditionally defective for undecylprodigiosin, and actinorhodin production is merely delayed in *bldB* mutants. *bldC, bldF* and *bldK* mutants are defective in production of at least one antibiotic. Additional *bld* mutants have been recently identified and mapped to unique loci on the *S. coelicolor* chromosome (Chater, 1993) (Redenbach *et al.*, 1996). These include *bld-261* (Willey *et al.*, 1993), *bld-221* (Willey *et al.*, 1991), and *bld-5M1* (Schauer *et al.*, 1991). *bld-221* is reported to be defective for antibiotic production and conditionally defective for aerial mycelia whereas *bld-5M1* is defective for both antibiotic and aerial hyphae production. Another *bld* mutant, *bld-5M5* could be an allele of *bldF* as it maps to the same region, although its phenotype is distinct (Schauer *et al.*, 1991). Most of the *bld* mutants will sporulate if grown in close proximity to a wild-type colony, indicating that most *bld* mutants may be defective in production of extracellular signalling molecules. Likewise, when two different *bld* mutants are grown in close proximity, sporulation is often restored to one of the mutants (Willey *et al.*, 1993), which likely indicates that exogenous addition of an absent signalling molecule can overcome the defect caused by most of the *bld* mutations. This extracellular complementation can be defined by a hierarchical cascade (*bld261 < bldK < bldA, bldH < bldG < bldC < bldD)* (Nodwell *et al.*, 1996) where *bldD* can restore sporulation to all the other *bld* mutants but cannot be rescued by any of them. Recently, Nodwell *et al.* (1999) has characterized additional *bld* mutants by extracellular complementation. Most of these mutants fall into the complementation groups K, D, and C, which are represented by *bldK, bldD,* and *bldC* respectively. Mapping of the new *bld* mutants indicates that while some of the mutations represent additional alleles of *bldK, bldD,* and *bldC,* some of the mutations map to new loci and likely represent at least three additional *bld* genes. The hierarchy of extracellular complementation of the *bld* mutants suggests that there are at least five extracellular signalling molecules regulating sporulation. One of the potential signal molecules has been recently purified from conditioned medium (Nodwell and

Losick, 1998), and response to this signal molecule is dependent on the BldK oligopeptide import system, indicating that this peptide is likely the first signal molecule in the hierarchy. However, some of the *bld* mutants *(bldB* and *bldI* as well as six *bld* mutants identified by Nodwell *et al.* (1999)), show an unusual pattern of extracellular complementation and cannot be placed into the hierarchy (Nodwell *et al.*, 1999; Willey *et al.*, 1993), indicating that the signals regulating sporulation and antibiotic production are more complex than a linear cascade of extracellular signalling molecules.

There are additional genes that have been implicated in the sporulation process in *S. coelicolor* that are not classified as either *whi* or *bld.* The gene encoding SapA, a 13 kilodalton spore-associated protein, has been cloned from *S. coelicolor* (Guijarro *et al.*, 1988). *sapA* was found to be expressed only in the aerial hyphae, and the level of transcript was significantly reduced in *bldC, bldD,* and *whiH* mutants. Another spore-associated protein, SapB, is not produced in any of the *bld* mutants but is produced normally in the *whi* mutants (Willey *et al.*, 1991). Growth of a SapB producing colony near a *bld* colony, or the addition of purified SapB directly to *bld* mutant colonies, can partially restore sporulation. Interestingly, SapB is not required for sporulation under all conditions, as both *bld* and wild-type strains grown on minimal medium with mannitol as carbon source sporulate without producing SapB. Recent studies (Tillotson *et al.*, 1998) indicate that SapB is a surfactant capable of reducing the surface tension of water, and aids in the formation of aerial hyphae by decreasing the surface tension at the colony surface. This suggests that in *bld* mutants SapB suppresses the deficiency in formation of erect aerial hyphae, perhaps by compensating for the lack of turgor pressure. At any rate, it appears that SapB only partially complements the *bld* phenotype, as SapB treated *bld* mutants are still defective for antibiotic production and maturation of aerial hyphae into mature spore chains, indicating that SapB cannot compensate for the absence of the *bld* gene products.

The *ram* (*r*apid *a*erial *m*ycelium) cluster consists of three genes, *ramA*, *ramB*, and *ramR* (Ma and Kendall, 1994). RamA and RamB show homology to ATP-dependent membrane-translocating proteins, and are predicted to form a heterodimer. RamR shows homology to the UhpA subset of two-component response regulator proteins that are known to play a role in sugar-phosphate transport in *E. coli* (Stock *et al.*, 1989; Weston and Kadner, 1988). In prokaryotic two-component systems, the sensor kinase and response regulator genes typically map to the same cluster, however, the corresponding sensor kinase that phosphorylates the RamR response regulator is not located in the *ram* cluster. This is similar to the situation with Spo0A of *B. subtilis* (Ferrari *et al.*, 1985; Hoch, 1993), which is the final receptor in a complex phosphorelay regulating sporulation, which suggests that RamR is part of a much larger network of response regulators and sensor kinases that controls development. A *ramB* disruption mutant was severely defective in aerial mycelium formation, indicating that this cluster is required for aerial mycelium formation and development.

The *amfRAB* (*a*erial *m*ycelium *f*ormation) gene cluster, that was found to restore sporulation to *S. griseus* A-factor mutants (*bld* phenotype) (Ueda *et al.*, 1993), shows homology at the amino acid level to the *ramRAB* cluster. *amfC* from *S. griseus* was also found to restore sporulation to A-factor mutants, and *amfC* disruption mutants were found to sporulate poorly (Kudo *et al.*, 1995). The *amfC* homologue was identified in *S. coelicolor*, and the *S. coelicolor amfC* gene complemented the *S. griseus amfC* mutant phenotype, suggesting that these two proteins perform similar functions in the two species. *nrsA* (*n*egative *r*egulator of *s*porulation, previously *orf1590*) is another sporulation gene that was first identified in *S. griseus* (Babcock and Kendrick, 1990; McCue *et al.*, 1996) . The gene was initially identified by virtue of its ability to restore sporulation and antibiotic production to certain *bld* mutants. *nrsA* null mutations can only be generated in strains containing additional mutations that block sporulation at an early stage, suggesting that deletion of this gene in a wild-type strain is lethal, and that the *nrsA* gene functions to prevent premature sporulation. Interestingly, the *nrsA* gene encodes two proteins from nested coding sequences. The larger protein, called P56, is translated from the first initiation codon and is believed to be responsible for suppression of the *bld* phenotype. It is postulated that the smaller protein, P49.5, which is translated from an alternate initiation codon (ATG55 of P56) modulates the activity of P56 and prevents premature sporulation. The *nrsA* counterpart, *bldX*, has been identified in *S. coelicolor* (McCue *et al.*, 1992), and the similarly of the deduced protein sequences suggests that BldX performs a similar function in regulating sporulation in *S. coelicolor*. The overall similarity in structure and function between the *amfRAB* and *ramRAB* clusters, *amfC* homologues, *nrsA* and *bldX*, and *bldA* homologues (McCue *et al.*, 1992) of *S. griseus* and *S. coelicolor* suggest that regulation of sporulation (and perhaps secondary metabolism) is similar in these two species and likely will prove to be similar in all *Streptomyces* species.

As mentioned previously, the hierarchy of extracellular complementation of the *bld* mutants suggests that there are at least five extracellular signalling molecules regulating sporulation. Although at present, only one potential signalling molecule has been identified in *S. coelicolor* (Nodwell and Losick, 1998) it is likely that at least one of the unidentified signalling molecules will prove to be a hormone of the γ-butyrolactone type as these molecules have been identified in numerous other *Streptomyces* species and are implicated in the regulation of differentiation and antibiotic production. Butyrolactones implicated in antibiotic production include IM-2 from *Streptomyces* sp. FRI-5 (Hashimoto *et al.*, 1992) and virginiae butanolides A-E from *Streptomyces virginiae* (Nihira *et al.*, 1988; Okamoto *et al.*, 1995). A-factor (2S-isocapryloyl-3S-hydroxymethyl-γ-butyrolactone) in *S. griseus* is implicated in both aerial mycelium formation and antibiotic production (Hara *et al.*, 1983), and Factor I is required for differentiation and anthracycline biosynthesis in *S. bikiniensis, S. cyaneofuscatus* and *S. viridochromogenes* (Grafe *et al.*, 1983).

S. griseus A-factor deficient mutants are defective in aerial mycelium formation, streptomycin production and resistance, and production of a yellow pigment (Miyake *et al.*, 1990). Although defective in aerial mycelium formation, and characterized as *bld,* A-factor mutants do in fact sporulate (Szabo and Vitalis, 1992). The spores are produced in the substrate mycelium and are more heterogeneous in size and shape than those produced on aerial hyphae, however, it is apparent that the defect imposed by the absence of A-factor is in the formation of aerial hyphae and not the production of spores. The mutant phenotype can be reversed by the exogenous addition of small quantities (10^{-9} M) of purified A-factor to the mutant colonies (Horinouchi and Beppu, 1990a), or by growing mutant colonies in close proximity to an A-factor producing strain.

A-factor is synthesized by the product of the *afsA* gene (Horinouchi, 1986), and exerts its effect through its interaction with its specific receptor protein, ArpA (*A*-factor receptor protein) (Onaka *et al.*, 1995). *afsAarpA* double mutants defective in the production of both A-factor and the A-factor receptor protein sporulate and produce streptomycin earlier in development than wild-type strains, and produce 10-fold more streptomycin than wild-type strains (Miyake *et al.*, 1990). This indicates that ArpA is a repressor type protein that negatively regulates differentiation and antibiotic production, and that A-factor functions to regulate the timing and level of expression of the ArpA repressed genes. The ArpA protein consists of two functional domains, a DNA binding domain and an A-factor binding domain (Onaka and Horinouchi, 1997b). In the absence of A-factor, homodimers of ArpA bind to a 22 bp palindromic sequence, and repress expression. Amino acid substitutions in the DNA binding domain cause early sporulation and antibiotic production indicating that the DNA binding activity of ArpA is required for repression of these two pathways (Onaka *et al.*, 1997a). In the presence of A-factor, ArpA is unable to bind DNA, suggesting that A-factor releases the repression imposed by ArpA by binding to the protein and inducing a conformational change that removes ArpA from the DNA binding site.

Although none of the genes directly regulated by ArpA have been identified, the *strR* (regulator of streptomycin production) promoter is A-factor dependent (Vujaklija *et al.*, 1991), as is the *amfR* transcriptional unit (Ueda *et al.*, 1998). Several proteins have been detected that bind to the *strR* promoter region, however, the one identified as being A-factor dependent is required for activation of *strR* transcription which indicates that this protein is not ArpA (Vujaklija *et al.*, 1993). However, this A-factor dependent protein (protein X) is a good candidate for repression by ArpA. Similarly, the *orf5-orf4-amfR* operon is A-factor dependent, however, it requires A-factor to relieve the repression imposed by AdpB (*A*-factor *d*ependent *p*rotein) (Ueda *et al.*, 1998). Although AdpB functions as a repressor and is responsive to A-factor, this protein is not ArpA as it has a different molecular weight and DNA binding recognition sequence. At present it is not known if A-factor interacts directly with the AdpB protein or if its effect is indirect.

A-factor-like substances that are able to restore streptomycin production to *S. griseus* A-factor mutants are produced in several other *Streptomyces* species (approximately 15% of *Streptomyces* species) as well as in a few *Actinomyces* species and one *Nocardia* species (Hara and Beppu, 1982), which indicates that A-factor or other γ-butyrolactones are likely present in all Actinomycetes and these effector molecules most likely play a role in differentiation and antibiotic production. An A-factor-like substance is required for nosiheptide production in *Streptomyces actuosus* and anthracycline production in *S. griseus* (Horinouchi and Beppu, 1992b). However, while an A-factor-like substance produced by *S. coelicolor* and *S. lividans* is able to restore streptomycin production to *S. griseus afsA* mutants, this substance is not a candidate for one of the unidentified signalling molecules required for differentiation as *S. coelicolor* mutants unable to synthesize this compound and therefore unable to complement *S. griseus* A-factor mutants do not display morphological or physiological defects (Hara *et al.*, 1983). However, *S. coelicolor* produces several other A-factor-like substances (Bibb, 1996), and while these substances are unable to complement *S. griseus* A-factor mutants it appears that at least one of them regulates aerial mycelium development and antibiotic production in *S. coelicolor*, similar to the role A-factor plays in *S. griseus*. Two ArpA homologues have recently been identified in *S. coelicolor* (Onaka *et al.*, 1998). These homologues are called *cprA* and *cprB* (*c*oelicolor *p*igment *r*egulator) and both of these genes are implicated in the regulation of actinorhodin and undecylprodigiosin production as well as aerial mycelium formation. CprB appears to negatively regulate actinorhodin production and aerial mycelium formation as *cprB* mutants sporulate and produce actinorhodin earlier than the wild-type parental strain. The similarities between CprB and ArpA in both amino acid sequence homology and phenotype of the respective mutants suggest that CprB has an analogous function to ArpA, and that CprB activity is modulated by one of the A-factor-like substances. In contrast, CprA appears to positively regulate antibiotic production and aerial mycelium formation as *cprA* mutants show delayed sporulation and produce less antibiotic than wild-type. CprA and CprB show 90% identity in amino acid sequence and could bind to the same DNA sequences as the two putative DNA binding domains are nearly identical. It remains to be determined if CprA and CprB interact with the same or different A-factor-like molecule, and whether these two proteins regulate the expression of the same set of genes.

While at least one, and possibly more, of the unidentified signalling molecules is likely to be an A-factor-like substance, it is also possible that others may be hormones that are not γ-butyrolactones. Additional substances that appear to function as bacterial hormones but are not γ-butyrolactones have been identified in *Streptomyces*. These include: factor C from *S. griseus* (Biro *et al.*, 1980); pamamycin from *Streptomyces alboniger* (Kondo *et al.*, 1988); sporulation pigment from *Streptomyces venezuelae* (Horinouchi and Beppu, 1992b); borrelidin from *Streptomyces tendae* (Schuz and Zahner, 1993), and a trypsin-like protease in *Streptomyces exfoliatus* (Kim and Lee, 1996). Calcium ions are also implicated in differentiation in a number of *Streptomyces* strains (Natsume *et al.*, 1989), and are also required for the production of one of the antibiotics in *S. coelicolor* (Hopwood and Wright, 1983; Lakey *et al.*, 1983), although the mechanism by which calcium ions induce aerial mycelium formation and antibiotic biosynthesis remains to be elucidated.

Recent studies indicate that *bld* mutants *(bldA, bldB, bldC, bldD, bldG,* and *bldH)* are also defective in the regulation of carbon utilization (Pope *et al.*, 1996), suggesting that some of the morphological and physiological defects imposed by some of the *bld* mutations may be the result of the inability of *bld* mutants to sense and/or signal starvation or other stress conditions, rather than defects per se in regulation of sporogenesis and secondary metabolite induction. It was once generally assumed that the onset of sporulation and secondary metabolism was triggered by nutrient limitation, however, Granozzi and co-workers (1990) found that transferring *S. coelicolor* cultures growing on cellophane disks to fresh medium did not delay the onset of aerial mycelium formation, suggesting that stress conditions other than starvation trigger aerial mycelium development.

One of the metabolites that may play a role in the initiation of secondary metabolism and morphogenesis is cAMP (cyclic adenosine 3',5'-monophosphate). *S. coelicolor* mutants unable to synthesize cAMP (*cya*- mutants) display a *bld* phenotype on unbuffered medium, and this phenotype is suppressed by exogenous addition of cAMP up until 20 hours post-germination (Susstrunk *et al.*, 1998). These observations, along with the observation that wild-type colonies accumulated cAMP just prior to the initiation of aerial mycelium formation and secondary metabolism, suggest that cAMP is required to trigger one or both of these processes. Additional tests indicate that *cya* mutants are unable to switch to an alternate (neutralizing) metabolic pathway after vegetative growth ceases. The consequence of this defect on unbuffered media is that the *cya* mutants continue to produce acidic end-products which increasingly acidify the surrounding medium until growth ceases when the pH drops below 4.5. Aerial mycelium formation and actinorhodin production can be restored by growth of *cya* mutants on buffered medium. Similarly *cya* mutants can sporulate and produce antibiotics along pH gradients created by addition of discs containing buffer or growth of wild-type colonies near *cya* mutants. Taken together, the results indicate that the irreversible acidification of the media by *cya* mutants is responsible for the *bld* phenotype.

While several of the *bld* mutants (*bldA, bldB, bldC, bldD* and *bldG)* irreversibly acidified the growth medium when grown on unbuffered media, they were not defective in their ability to produce cAMP, and the *bld* phenotype was not rescued by exogenous addition of cAMP or growth on buffered medium (Susstrunk *et al.*, 1998). So while it is clear that under some conditions the defect in carbon catabolite repression of *bld* mutants results in growth cessation and a *bld* phenotype due to excess acid production, it is not clear if the defect in carbon utilization is responsible for the *bld* phenotype under all conditions.

Additional stress conditions that may trigger antibiotic production and sporulation could be mediated by the stringent response. Ochi (1990a) has proposed that (p)ppGpp (guanosine 5'-triphosphate 3'-diphosphate and guanosine 5'-diphosphate 3'-diphosphate) accumulation induces antibiotic production and a decrease in GTP pool induces morphological differentiation. This hypothesis was based largely on studies of stringent response induction and analysis of *rel* (relaxed) mutants of *S. griseus* (Ochi, 1987a; Ochi, 1987b; Ochi, 1990b), and may not reflect the stringent response in *S. coelicolor*. However, studies of *rel* mutants in *S. coelicolor* indicate that the stringent response likely functions in a similar manner to that observed in *S. griseus*. *S. coelicolor relC* mutants lacking the ribosomal protein ST-L11 which interacts with the (p)ppGpp synthetase I enzyme display severely delayed sporulation and actinorhodin production and are defective for undecylprodigiosin production (Ochi, 1990a). In contrast, *relA* mutants defective in the (p)ppGpp synthetase I enzyme are not defective in antibiotic production or sporulation, although the growth rate is reduced and isolated colonies are smaller than wild-type. However, disruption of another gene that encodes a (p)ppGpp synthetase enzyme that shows homology to both *relA* and *spoT* synthetases, delays sporulation and disrupts actinorhodin production (Martinez-Costa *et al.*, 1996). While these results are difficult to interpret, it appears that the stringent response does play a role in both differentiation and antibiotic production. Whether the defect in carbon utilization of the *bld* mutants has any effect on the stringent response remains to be determined. It seems possible that the metabolic defect could result in altered or unbalanced nucleotide pools which could affect the production of (p)ppGpp and ultimately affect the induction of sporulation and antibiotic production.

While it is unclear what role carbon catabolite repression and the stringent response play in morphological and physiological differentiation, it is becoming increasing apparent that sporogenesis and secondary metabolite production are regulated by signal transduction as protein phosphorylation and two-component phosphorelay systems have been implicated in these processes (Horinouchi, 1993). The *whiI* locus consists of a sensor-kinase and response regulator implicated in aerial mycelium formation (K. Chater, John Innes Centre http://www.jic.bbsrc.ac.uk/). The *ramRAB* locus in *S. coelicolor* (Ma and Kendall, 1994), and the *S. griseus* homologue *amfRAB* (Ueda *et al.*, 1993) encode ATP-dependent membrane-bound proteins involved in translocation (RamA, RamB and AmfA, AmfB) and response-regulator proteins (RamR and AmfR) which are implicated in aerial mycelium formation and sporulation. *afsR/afsK* (Horinouchi, 1986; Matsumoto *et al.*, 1994) and *afsQ1/afsQ2*(Ishizuka *et al.*, 1992) encode two-component regulatory systems *(afsR/afsQ1* encode response regulators and *afsK/afsQ2* encode sensor kinases) that are implicated in secondary metabolism in *S. coelicolor*. Since AfsR is phosphorylated to some extent in *afsK* mutants (Matsumoto *et al.*, 1994), AfsR may also be phosphorylated by an additional, as yet unidentified, sensor-kinase. AfsR/AfsK show characteristics of eukaryotic two-component regulators, which consist of Ser/Thr phospho-transfer systems (Matsumoto *et al.*, 1994), whereas the AfsQ1/AfsQ2 proteins are typical prokaryotic Asp/His two-component regulators (Horinouchi, 1993). Homologues of *afsQ1/afsQ2* have been found in almost all *Streptomyces* species tested (Ishizuka *et al.*, 1992) and homologues of *afsR/afsK* are also prevalent throughout the genus (Horinouchi, 1993). There is substantial evidence to suggest that there are many additional two-component regulators (Hirakata *et al.*, 1998; Hong *et al.*, 1993; Urabe and Ogawara, 1995; Waters *et al.*, 1994), phosphatases (Li and Strohl, 1996; Umeyama *et al.*, 1996), ADP-ribosylated proteins (Penyige *et al.*, 1990; Penyige *et al.*, 1992; Shima *et al.*, 1996), GTP-binding proteins (Itoh *et al.*, 1996; Okamoto *et al.*, 1997; Okamoto and Ochi, 1998), as well as additional phosphorylated proteins (Okamoto *et al.*, 1998; Smardova *et al.*, 1992), and these proteins are implicated in morphogenesis or antibiotic production. The existence of a complex signal transduction system regulating differentiation indicates that nucleotide pools, primarily GTP and ATP pools, may provide the link between primary metabolism, carbon catabolite control, the stringent response and morphological and physiological differentiation in *Streptomyces.*

Several genes have been identified that are required for sporulation but are also implicated in cell-division in vegetative mycelium. The first gene *ftsZ* was isolated from both *S. coelicolor* (McCormick *et al.*, 1994) and *S. griseus* (Dharmatilake and Kendrick, 1994). As the name implies, these proteins are homologues of the FtsZ protein from *E. coli* and other bacteria. In *E. coli,* the FtsZ protein is absolutely required for cell-division and viability (Dai and Lutkenhaus, 1991) and in *B. subtilis,* FtsZ is required for septum formation, cell division and endospore formation (Beall and Lutkenhaus, 1991). This differs somewhat from the situation in *Streptomyces*, where the *ftsZ* null mutant is viable (McCormick *et al.*, 1994) but is clearly aberrant in septum formation, in both vegetative mycelia and aerial hyphae, which indicates that this gene functions during both vegetative growth and differentiation. The *ftsQ* gene of *S. coelicolor* is similarly required for cross-wall and sporulation septa formation but not for growth or viability (McCormick and Losick, 1996). The *ssgA* gene is implicated in both sporulation and cell division in *S. griseus,* as the gene in high copy suppresses sporulation and causes fragmentation of mycelia (Kawamoto *et al.*, 1997).

Despite all of the research that has been aimed at understanding the process of morphological differentiation, *Streptomyces* are still best known for their ability to synthesize a vast array of complex secondary metabolites. Many of these compounds have therapeutic applications as antibacterial, antitumor, or antifungal agents (Bibb, 1996). Other compounds have applications in the agricultural industry as growth promoters, fungicides, herbicides, or antiparasitic agents (Bibb, 1996). There are ten major structural classes of antibiotics produced by streptomycetes (Crandall and Hamill, 1986) including: aminocyclitols, ansamycins, anthracyclines and other quinones, β-lactams, macrolides, nucleosides, peptides, polyenes, polyether antibiotics, and tetracyclines. Although the chemical structures of the different classes of antibiotics and their biosynthetic genes may share no similarities, the overall organization and regulation of the biosynthetic genes for antibiotic production share some common features. The antibiotic biosynthetic genes are clustered on the chromosome along with regulatory elements and genes required for export and/or resistance. Tandemly arranged genes are often transcribed as a polycistronic message, and the cotranscribed genes typically encode enzymes required for sequential biosynthetic steps. Moreover, the genes required for the initial biosynthetic steps are typically grouped as are the genes required for late biosynthetic steps. The clustering of the biosynthetic genes along with resistance and regulatory genes makes the study of antibiotic production convenient and amenable to genetic manipulation.

As stated previously, the genetically most-studied streptomycete, *Streptomyces coelicolor* produces four antibiotics: actinorhodin (Wright and Hopwood, 1976b), undecylprodigiosin (Rudd and Hopwood, 1980), methylenomycin (Wright and Hopwood, 1976a), and a calcium-dependent antibiotic (Hopwood and Wright, 1983). Although none of these antibiotics have therapeutic value, actinorhodin and undecylprodigiosin are pigmented and are useful for the study of antibiotic production as the colored products can be easily detected visually or quantified spectrometrically.

Actinorhodin is an intensely blue-colored pigment that diffuses into the media surrounding the colonies (Wright and Hopwood, 1976b). It belongs to the class of antibiotics that include anthracyclines and other quinones along with daunorubicin (daunomycin) and doxorubicin (adriamycin) produced by *Streptomyces peucetius* (Arcamone *et al.*, 1969; Di Marco *et al.*, 1964) which are used in cancer therapy. Members of this group of antibiotics, as well as the macrolide class of antibiotics are also commonly referred to as polyketide antibiotics because they are synthesized through a series of condensations of carbon units in a manner similar to that of long chain fatty acid biosynthesis (Hopwood and Sherman, 1990). The pathway specific activator and export genes for the 25 kb actinorhodin biosynthetic cluster are encoded by the *act*II region (Fernández-Moreno *et al.*, 1991). *act*II-ORF1 encodes a putative repressor protein that regulates the transcription of the *act*II-ORF2 and *act*II-ORF3 genes which in turn encode proteins that form the antibiotic export complex (Caballero *et al.*, 1991). The last gene in the *act*II region (*act*II-ORF4) encodes the pathway-specific activator which is required for activation of the actinorhodin biosynthetic genes encoded by the *act*I, III, IV, V, VI and VII regions (Malpartida and Hopwood, 1984).

Although the red, cell-associated pigment produced by *S. coelicolor* is usually referred to as undecylprodigiosin, the pigment actually consists of a mixture of tripyrrole prodigiosin derivatives (undecylprodigiosin, butylcycloheptylprodiginine, and to a lesser extent dipyrrolyldipyrromethane) (Tsao *et al.*, 1985). The complete biosynthetic cluster for undecylprodigiosin (red) biosynthesis is approximately 36 kb (Malpartida *et al.*, 1990) and is regulated by two pathway-specific regulators, RedD (Guthrie and Chater, 1990) and RedZ (White and Bibb, 1997). The *redZ* gene product is required for activation of the *redD* gene, and RedD is required for activation of the *red* biosynthetic genes (White and Bibb, 1997). The RedD transcriptional activator shows homology to other known pathway-specific activators such as ActII-ORF4, while the RedZ activator resembles typical two-component system response-regulators. However, RedZ lacks several highly conserved amino acid residues that form the phosphorylation pocket, including the aspartate residue that is phosphorylated by the sensor kinase (Guthrie *et al.*, 1998).

The biosynthesis of methylenomycin has not been extensively studied, however, it appears that the methylenomycin product or one of its intermediates induces transcription of the methylenomycin resistance gene (Hobbs *et al.*, 1992). This is somewhat unusual as typically the antibiotic resistance determinant is produced before the antibiotic. The methylenomycin biosynthetic gene cluster is located on the SCP1 plasmid (Wright and Hopwood, 1976a), and is the only example thus far in *Streptomyces* of an entire antibiotic cluster mapping to an extrachromosomal element (Hopwood, 1978). The fourth antibiotic produced by *S. coelicolor* is a calcium-dependent antibiotic (Hopwood and Wright, 1983; Lakey *et al.*, 1983). Since the 35 kb cluster has only recently been cloned (Chong *et al.*, 1998), little is known about the biosynthetic genes or their regulation.

Although much remains to be elucidated about the regulation of antibiotic biosynthesis and the regulation of secondary metabolites in general, it is apparent that the biosynthesis of these compounds is subject to several levels of regulation. Most antibiotic pathways are regulated by one or more pathway-specific regulators, such as ActII-ORF4, RedZ and RedD. There are also global regulatory factors that influence both antibiotic production and cellular differentiation such as the *bld* gene products and bacterial hormones. There is another level of regulation for antibiotic production that acts after the global regulators but before the pathway-specific regulators. None of these regulators are implicated in morphological development but most are pleiotropic in that they regulate the production of more than one antibiotic or secondary metabolite in the same organism. Several such regulatory genes or loci have been identified in *S. coelicolor* and include: *abaA* (Fernández-Moreno *et al.*, 1992), *abaB* (Scheu *et al.*, 1997), *afsR/afsK* (Matsumoto *et al.*, 1994), *afsB* (Horinouchi *et al.*, 1983), *afsS* (Matsumoto *et al.*, 1995; Vogtli *et al.*, 1994), *afsQ1/afsQ2* (Ishizuka *et al.*, 1992), *absA* (Adamidis *et al.*, 1990), *absB* (Adamidis and Champness, 1992), and *mia* (Champness *et al.*, 1992). *afsB* mutants are defective in the production of pigmented antibiotics as well as the production of the A-factor-like substance that complements *S. griseus* A-factor deficient mutants (Horinouchi *et al.*, 1983). Although *afsB* mutants were among the first *afs* mutants isolated in *S. lividans*, the *afsB* gene has not yet been identified. Most of the other *afs* genes were originally isolated by virtue of their ability to complement the *afsB* mutant phenotype (*afsR/afsK, afsS,* and *afsQ1/afsQ2).* The antibiotic deficiency of *afsB* mutants is also complemented by *hrdB,* which encodes the major sigma factor of *S. coelicolor* (Shiina *et al.*, 1991). *afsR/afsK* and *afsQ1/afsQ2,* as previously discussed, encode two-component regulators that positively regulate the production of the two pigmented antibiotics, undecylprodigiosin and actinorhodin in *S. coelicolor,* as well as the production of one of the A-factor-like substances. One of the two open reading frames located between *afsR* and *afsK,* named *afsS* (Matsumoto *et al.*, 1995) or *afsR2* (Vogtli *et al.*, 1994) encodes a small protein that stimulates pigment production in *S. lividans,* when present in high copy number.

The *abaA* locus from *S. coelicolor* is capable of stimulating actinorhodin production in *S. lividans* (Fernández-Moreno *et al.*, 1992). The locus consists of five open-reading frames, *orfsA-E,* although *orfB* alone stimulates actinorhodin production in *S. lividans* and *orfB* disruption in *S. coelicolor* impairs actinorhodin, undecylprodigiosin, and calcium-dependent antibiotic production, indicating that *abaA-orfB* positively regulates the production of three of the four antibiotics in *S. coelicolor.* The *abaB* locus, from *S. antibioticus,* stimulates actinorhodin and undecylprodigiosin production in *S. lividans* (Scheu *et al.*, 1997). The *abaB* gene encodes a LysR-type transcriptional regulator, and Southern analysis indicates that this gene is highly conserved among the *Streptomyces* species. The *abaB* promoter region in high copy number was sufficient to stimulate antibiotic production, indicating that the promoter region might contain DNA-binding sites that in high copy titrate out a repressor or other effector protein to stimulate antibiotic production.

Two classes of *S. coelicolor* mutants, *absA* and *absB,* are defective in production of all four antibiotics (Adamidis *et al.*, 1990). The *absA* locus consists of a sensor histidine kinase and response regulator (Brian *et al.*, 1996), typical of eubacterial two-component signal transducers. Surprisingly, disruption of this locus did not result in loss of pigmentation as expected, but rather caused precocious overproduction of antibiotics, indicating that the *absA* locus negatively regulates antibiotic production. The nature of the *absA* mutants are unknown, but since they all map to the sensor kinase it seems likely that the mutations result in a defect in the ability of the sensor kinase to respond to external stimuli and relieve repression of antibiotic biosynthetic pathways. The *absB* locus encodes a homologue of *E. coli* RNase III (Aceti and Champness, 1998). How this enzyme is involved in the activation of antibiotic production remains to be determined although it could potentially exert its effect at the level of the pathway-specific regulators as the level of *act*II-orf4 and *redD* transcripts are reduced in *absB* mutants.

There is evidence for an additional locus that negatively regulates antibiotic production. This locus called *mia* (for *m*ulticopy *i*nhibition of *a*ntibiotic production) inhibits antibiotic production when present on a high-copy number vector but not when present on a low-copy number vector (Champness *et al*., 1992). The *mia* locus doesn't appear to encode a protein so additional tests will have to be performed to determine the nature of this sequence.

afsA and *afsB,* by strict definition are the only truly global regulators of antibiotic production as these loci regulate production of all four antibiotics produced by *S. coelicolor.* The *abaA* locus is implicated in the regulation of three of the four antibiotics and it is compelling to speculate that it represents an additional level of regulation followed by the Afs class of regulators. The ability of cloned regions of the *afsR/afsK/afsS* locus to restore pigmented antibiotic production, but not methylenomycin or calcium dependent antibiotic production, to both *absA* and *absB* mutants (Champness *et al*., 1992) supports this hypothesis since only partial complementation would be expected if the *afsR/afsK/afsS* genes function after the *abs* genes. There may be yet an additional level of regulation represented by the *cutRS* signal transduction system of *S. lividans* (Chang *et al*., 1996). The *cutRS* genes encode a histidine protein kinase and its response regulator which negatively regulate actinorhodin production in both *S. lividans* and the closely related *S. coelicolor.* It is currently not known whether the *cutRS* signal transduction system regulates the production of any other antibiotics in either *S. lividans* or *S. coelicolor* although the locus is most likely pleiotropic as it is thought to also regulate copper metabolism (Tseng and Chen, 1991). Whether these regulators of antibiotic biosynthesis function in an hierarchical cascade, similar to the cascade observed with the various *bld* and *whi* genes remains to be determined, however, since these genes exert their effect at a specific point in development it is likely that they interact, either directly or indirectly. It is also likely that most of the global regulators of antibiotic production in *S. coelicolor* have homologues in other *Streptomyces* species, since the *abaA*, *abaB,* and most of the *afs* loci hybridize to chromosomal DNA from other *Steptomyces* species.

Overall, morphogenesis and antibiotic production appear to be regulated by complicated cascades of regulatory elements, including phosphorelay systems, sigma factors, transcription factors and small effector molecules that function as microbial hormones. These effectors function at the end of vegetative growth and their production appears to be influenced by products of primary metabolism, such as cAMP, as well as by the stringent response. While there are numerous examples of regulatory elements, most of these ultimately exert their effect at the level of transcription. The action of *bldA* is unusual in this respect since the *bldA* gene encodes a tRNA that translates the rare leucine codon UUA (Lawlor *et al.*, 1987) and is predicted to exert its effect at the level of translation. *Streptomyces* DNA has an extremely high G+C content (~70%), and in order to achieve this high G+C content, codon usage is biased to contain G or C in the third codon (wobble) position, and to a lesser extent in the first codon position (Bibb *et al.*, 1984). The result is that some codons such as the UUA leucine and UUU phenylalanine codons are used much less frequently than CUC or CUG leucine and UUC phenylalanine codons, and are considered rare or minor codons in *Streptomyces* mRNA. The codon usage in *Streptomyces* varies from the most rare UUA codon (used at a frequency of 0.2/1000 codons) to the most common GCC alanine codon (used at a frequency of 76.1/1000 codons) (*Streptomyces* Internet Resource Center, http://molbio.cbs.umn.edu/asirc/lib/lib.html). While there is no question that the UUA leucine codon is minor or rare, there is no strict definition in the literature that indicates at what frequency a codon is considered minor or rare. While the GGU codon, used at a frequency of 11.1/1000 codons, is certainly a minor glycine codon in *Streptomyces,* as the major glycine codon, GGC, is encoded at a frequency of 58.9/1000 codons, it is not really a minor codon when compared to other codons, such as UUA (leucine) at 0.2/1000 codons, or CCU (proline) at 1.8/1000 codons. Since a definition of rare and minor codons was required for data analysis, I have taken the liberty of defining minor *Streptomyces* codons as codons that occur at a frequency of 5/1000 codons or less, and rare codons as codons that occur at a frequency of <1/1000 codons (Appendix 1). By this definition, there are six *Streptomyces* codons that are minor and rare: UUU (Phe), UUA (Leu), CUA (Leu), AUA (Ile), UCU (Ser), and AGA (Arg). The remainder of codons ending in A and U are minor with the exception of GCA (Ala), GAA (Glu), CGU (Arg), GGU (Gly), and GGA (Gly).

There is a positive correlation between the frequency of codon usage and the level of cognate charged-tRNA (Bulmer, 1987; Ikemura, 1981a). Major codons are translated by tRNAs expressed at high levels (major tRNAs) and minor codons are translated by tRNAs expressed at low levels (minor tRNAs). Typically minor codons are randomly distributed throughout the genome in genes expressed at various stages of growth.

Although there is a tendency for minor codons to be excluded from highly expressed genes, there is no proposed regulatory function for the cognate tRNAs required to translate the majority of these minor codons, and the cognate tRNAs are typically present at low levels throughout growth. In contrast, some minor tRNAs appear to have a regulatory function, as the minor codons they translate are confined to a particular subset of genes, expressed at a particular point in growth. Examples of minor tRNAs that have a proposed regulatory function are the *argU* (Garcia *et al.*, 1986) and *leuX* tRNAs (Newman *et al.*, 1994) of *Escherichia coli,* the *thrA* tRNA of *Clostridium acetobutylicum* (Sauer and Dürre, 1992), and the *bldA* tRNA of *Streptomyces coelicolor.* The *argU* tRNA translates the rarest codons in *E. coli,* AGA and AGG (Spanjaard *et al.*, 1990), and the tRNA is implicated in the regulation of stationary phase gene expression (Chen and Inouye, 1994). Genes containing AGA/AGG codons, such as several genes required for cell division and DNA replication are only translated efficiently during exponential growth when the *argU* tRNA is present. The *leuX* tRNA translates the rare leucine codon UUG, and regulates type 1 fimbrial expression in uropathogenic *E. coli* strains by modulating the expression of the *fimB* recombinase gene, containing five UUG codons (Ritter *et al.*, 1997). The *thrA* tRNA translates the rare threonine ACG codon in *C. acetobutylicum* and is implicated in the regulation of solventogenesis (Sauer and Dürre, 1992), and the *bldA* tRNA of *S. coelicolor* is implicated in the regulation of sporulation and antibiotic production (Lawlor, 1987; Leskiw *et al.*, 1993). While most tRNA genes in *E. coli,* and to a lesser extent in *Streptomyces,* are clustered on the chromosome and co-ordinately regulated, the *argU, leuX* and *bldA* genes are not part of a cluster and are not transcribed together with other tRNAs. The most likely reason for their isolation is that they are either highly regulated or uniquely regulated. Evidence in support of this statement comes from a study of Rowley et al. (1993) which found that four operons containing leucyl-tRNAs (*leuX, leuV, metT,* and *argT)* responded differently to leucine starvation, growth rate-dependence, analog inhibition, and stringent response induction, and that the *leuX* gene in particular appeared to be subject to multiple levels of control.

In addition to the non-random distribution of some rare codons with respect to the type of genes they are located in, some rare codons also show a non-random distribution within the target gene. In *E. coli,* the minor codons AGA and AGG are located preferentially within the first 25 codons, particularly if only one AGA or AGG codon is encoded in the target gene (Chen and Inouye, 1990). Translation of AGA and AGG codons is inefficient during stationary phase if the AGA/AGG codons are located within the first 25 codons, and as the distance from the start of the gene to the AGA/AGG codons increases so does the efficiency of translation. These observations lead Chen and Inouye to propose that the position of minor codons in an mRNA regulates gene expression by modulating the stability of the initiation complex for protein synthesis. According to this "minor codon modulator hypothesis", ribosomes translating mRNA pause when they reach a minor codon until a charged cognate tRNA becomes available. The paused ribosome-mRNA complex either becomes unstable and dissociates if it is near the initiation codon or the complex prevents additional ribosomes from initiating translation. The more limited the availability of charged minor tRNA the longer the pause. If this pause occurs at a distance greater than 50 or 60 codons from the initiation site, the paused complex either becomes stable and doesn't dissociate or the distance is sufficient to allow additional ribosomes to initiate translation (Chen and Inouye, 1994). Although this hypothesis was proposed to explain the mode of regulation of the *argU* tRNA at AGA/AGG codons, it may be generally applicable to global gene expression as other minor *E. coli* codons (CUA, UCA, AGU, ACA, GGA, CCC, and AUA) are preferentially located within the first 25 codons (Chen and Inouye, 1990). While the minor tRNAs that decode these minor codons do not appear to have a specialized function, it is possible that the availability of these tRNAs varies under certain conditions and that they do have an effect on the overall levels of protein synthesis. Alternately, minor codons, particularly near the start of genes might function to reduce the level of expression of weakly expressed genes, in contrast to highly expressed genes which lack these codons and are optimized for maximal expression (Folley and Yarus, 1989; Ikemura, 1981b). It is also likely that other minor tRNAs that do appear to have a specialized function, such as the *bldA* tRNA of *S. coelicolor,* modulate gene expression in a similar manner.

While the role of the *bldA* tRNA has only been studied extensively in *S. coelicolor,* it is likely to perform a similar function in all *Streptomyces* species as *S. griseus* and *S. lividans bldA* mutants display a similar phenotype and UUA codons are exceedingly rare throughout the genus. In *S. coelicolor,* the *bldA* tRNA has been shown to be the only tRNA capable of efficiently translating UUA codons since the TTA-containing reporter genes: *ampC, lacZ,* and *carB,* are not expressed in a *bldA* mutant (Leskiw *et al.*, 1991b). In addition, transcription analysis on surface-grown cultures of *S. coelicolor* has revealed that the appearance of the *bldA* tRNA is temporally regulated. Although the tRNA is transcribed throughout growth, the mature, processed form of the tRNA isn't detectable until the point in growth that coincides with the initiation of aerial mycelium formation, and gradually accumulates throughout the late growth phases (Leskiw *et al.*, 1993). Presumably the *bldA* primary transcripts are labile during early growth, as the mature, processed form of the tRNA is not detectable by northern analysis during this time and translation of UUA codons is inefficient (Leskiw *et al.*, 1991a).

While *bldA* mutants are defective in aerial mycelium formation and antibiotic production, they are capable of generating substrate mycelium, suggesting that UUA codons are confined to genes required for sporulation and antibiotic production and absent from genes expressed early in growth. However, the defect in carbon utilization and irreversible acidification of the surrounding medium under some conditions suggests that some vegetatively expressed genes contain UUA codons, and the inability of *bldA* mutants to translate these UUA-containing mRNAs causes this phenotype. A survey of all *Streptomyces* genes containing UUA codons indicates that UUA codons are primarily confined to genes expressed relatively late in growth such as: antibiotic resistance and export genes, antibiotic pathway-specific regulatory genes and biosynthetic genes, and genes encoding extracellular enzymes and transport proteins (Table I.1). In the *bldA* target sequences, 31 (29.2%) of the UUA codons are located within the first 25 codons and 47 (44.3%) are located within the first 50 codons, compared with only 15 (14.1%) of the UUA codons located within codons 51-100. Of the 96 UUA-containing genes (106 UUA codons) identified, 82 (85.4%) genes have more than 200 codons, and only one gene has fewer than 100 codons. Therefore, the high proportion of UUA codons in the first 50 codons is not indicative of random distribution in short protein sequences but reflects a preference for the location of these rare codons near the start of the gene. Since

Table I.1 ***Streptomyces bldA*** **targets**

Gene description*	**Organism and gene designation**	**# UUA codons**	**Position+ (codon)**
Antibiotic Resistance Genes:			
tylosin (MLS)[1]	*S. fradiae (tlrD)*	1	76 /327
streptomycin[2]	*S. glaucescens (strA)* formerly *(sph)*	1	39 /307
hygromycin[3]	*S. hygroscopicus (hyg)*	1	7 /332
kanamycin*[4]	*S. kanamyceticus (kmr)*	1	12 /276
mitomycin C*[5]	*S. lavendulae (mcr-* ORF3)	1	83 /281
midecamycin (MLS)[6]	*S. mycarofaciens (mdmA)*	3	59,63,85 /310+
carbomycin[7]	*S. thermotolerans (carB)*	2	95,111 /283
Antibiotic Export Genes:			
actinorhodin[8]	*S. coelicolor (act* II- ORF2)	1	19 /578
cephamycin and clavulanic acid*[9]	*S. clavuligerus (cmcT)*	1	470 /523
streptomycin[10]	*S. glaucescens (strV)*	1	4 /584
Antibiotic Regulatory Genes:			
spiramycin[11]	*S. ambofaciens (srmR)*	1	237 /604
secondary metabolism[12]	*S. antibioticus (abaB)*	1	122 /301
mithramycin[13]	*S. argillaceus (mtmR)*	2	7, 227 /276
actinorhodin[8]	*S. coelicolor (act* II-ORF4)	1	5 /255
undecylprodigiosin[14]	*S. coelicolor (redZ)*	1	156 /217
cephamycin and clavulanic acid[15]	*S. clavuligerus (ccaR)*	1	32 /256
spectinomycin*[16]	*S. flavopersicus (spcR)*	1	16 /330
streptomycin[2]	*S. glaucescens (strR)*	1	30 /424
streptomycin[17]	*S. griseus (strR)*	1	30 /350
bialaphos[18]	*S. hygroscopicus (brpA)*	1	250 /256
polyketide antibiotic*[19]	*S. hygroscopicus (orf6)*	1	369 /948
daunorubicin (repressor)*[20]	*S. peucetius (dnrO)*	1	86 /340
nogalamycin activator[21]	*S. nogalater (snoA)*	1	76 /665
Genes for differentiation:			
Negative regulator of sporulation[22]	*S. griseus (nrsA)* formerly *orf1590*	1	263 /529
Negative regulator of sporulation*[23]	*S. coelicolor (bldX* homologue of *S. griseus nrsA)*	1	207 /473
Aerial mycelium formation[24]	*S. griseus (amfR)*	1	141 /201
Regulatory Proteins:			
putative ATP/GTP binding protein[25]	*S. coelicolor* (SC1F2.20)	2	17,184 /731
Regulator of lipase[26]	*S. exfoliatus* M11 *(lipR)*	1	831 /934
Ser/Thr protein kinase[27]	*S. granaticolor (pkg4)*	1	114 /761
Regulator of cholesterol oxidase*[28]	*S. hygroscopicus (rapH)*	1	38 /872
DNA binding protein*[28]	*S. hygroscopicus (rapG)*	1	252 /330
Inhibitor gene for metalloproteinase[29]	*S. nigrescens* (SMPI)	1	7 /102
regulator of *acyB1* (carbomycin)*[30]	*S. thermotolerans (acyB2)*	1	228 /387
Antibiotic Biosynthesis Genes:			
polyketide synthase (oleandomycin)*[31]	*S. antibioticus* (OLE-ORF3)	2	242, 770 /3519
oxido-reductase (puromycin)*[32]	*S. anulatus (pur10)*	1	29 /338
puromycin synthetase*[32]	*S. anulatus (pur6)*	1	3 /772
polyketide synthase (niddamycin)[33]	*S. caelestis (nidA1)*	1	4004 /4340
polyketide synthase (niddamycin)[33]	*S. caelestis (nidA4)*	1	155 /1569
cytochrome P-450$_{sca\text{-}2}$ (pravastatin)[34]	*S. carbophilus* (ORF)	1	67 /410
glycosyltransferase (tylosin)*[35]	*S. fradiae (tyl-* ORF2 or *tylM2)*	1	264 /452
polyketide synthase (tylactone)*[36]	*S. fradiae (tylG-* ORF1)	1	3521 /4472
polyketide synthase (tylactone)*[36]	*S. fradiae (tylG-* ORF2)	1	1263 /1864
polyketide synthase (tylactone)*[36]	*S. fradiae (tylG-* ORF3)	1	1271 /3729
aminotransferase (N-methyl- L-glucosamine)[37]	*S. glaucescens (strS)*	1	7 /378
biosynthesis (streptomycin)[37]	*S. glaucescens (strD)*	1	14 /356

* Denotes putative function

+Position of TTA codon/Total number of codons

Targets were identified through codon usage tables at the on-line library at the Actinomycetes Internet Resouce Centre http://molbio.cbs.umn.edu/asirc/lib/lib.html.

Gene description	Organism and gene designation	# UUA codons	Position (codon)
Antibiotic Biosynthesis Genes:			
phosphotransferase (streptomycin)[37]	*S. glaucescens (strN)*	1	2 /315
O-methyltransferase (tetracenomycin)[38]	*S. glaucescens (tcmP)*	1	226 /271
apoprotein (C-1027 Antibiotic)[39]	*S. globisporus (axnA)*	1	3 /143
streptomycin-3'-phosphotransferase[40]	*S. griseus (orf* next to *aphE)*	1	47 /297
phosphotransferase or phosphatase[41]	*S. griseus (strN)*	2	2,36 /339
polyketide synthase*[42]	*S. hygroscopicus (orf4)*	1	655 /1937
polyketide synthase (rapamycin)[43]	*S. hygroscopicus (rapB)*	1	4578 /10223
pipecolate incorporation (rapamycin)[43]	*S. hygroscopicus (rapP)*	1	768 /1541
ACV synthetase (cephamycin)[44]	*S. lactamdurans (pcbAB)*	2	775, 3456 /3649
lincomycin*[45]	*S. lincolnesis (lmbB2)*	1	104 /317
lincomycin*[45]	*S. lincolnesis (lmbY)*	1	19 /295
lincomycin*[45]	*S. lincolnesis (lmbU)*	1	6 /223
acyltransferase[46]	*S. mycarofaciens (mdmB)*	1	20 /387
epimerase (nogalamycin)*[47]	*S. nogalater (snoG)*	1	328 /328
glycosyltransferase (daunomycin)[48]	*Streptomyces sp.* C5 *(dauH)*	1	269 /442
polyketide synthase (FK506)[49]	*Streptomyces sp. (fkbA)*	1	4185 /6420
cytochrome P-450 (carbomycin)*[50]	*S. thermotolerans* (ORFA)	1	105 /411
p-aminobenzoic acid synthase (chloramphenicol)[51]	*Stretpomyces venuzuelae (pabAB)*	1	1779 /1949
PTT-synthetase A[52]	*S. viridochromogenes (phsA)*	1	222 /622
Other Enzymes and Proteins:			
ribonuclease Sa3*[53]	*S. aureofaciens (rnaSa3)*	1	52 /181
extracellular protease*[54]	*S. coelicolor* (SC10A5.18)	1	4 /411
extracellular hydrolase*[55]	*S. coelicolor* (SC4H2.07)	1	8 /400
aminotransferase*[55]	*S. coelicolor* (SC4H2.20)	1	525 /532
cytochrome P-450$_{SU1}$ (herbicide metabolism)[56]	*S. griseolus (suaC)*	1	26 /406
membrane translocase[57]	*S. griseus (secY)*	1	248 /436
endoglucanase[58]	*S. halstedii (celA2)*	1	4 /377
tyrosinase (copper transfer)[59]	*S. lavendulae* (TYR1-ORF1)	1	4 /156
α-amylase[60]	*S. lividans (amy)*	1	731 /919
endoglucanase[61]	*S. lividans (celB)*	2	4, 22 /381
chitinase[62]	*S. plicatus (chtA)*	1	276 /610
endoglucanase[63]	*S. rochei (eglS)*	1	4 /382
cytochrome P-450 (cholesterol oxidation)*[64]	*Streptomyces sp.* SA-COO *(choP)*	1	56 /381
L-proline 3-hydroxylase (collagen)[65]	*Streptomyces sp.* TH1 *(orf3)*	1	36 /870
protein secretion*[66]	*S. virginiae (secE)*	1	17 /121
Plasmid Integration and Transfer:			
pSAM2 transfer[67]	*S. ambofaciens (traSA)*	1	46 /415
pSAM2 gene of unknown function[68]	*S. ambofaciens (orf183)*	1	12/183
pBL1 gene of unknown function[69]	*S. bambergiensis* (ORF4)	2	80, 83 /98
pBL1 gene of unknown function[69]	*S. bambergiensis* (ORF7)	1	3 /109
pSC1 attL[70]	*S. griseus (attL)*	1	635 /666
traSA:integrase fusion protein[71]	*S. coelicolor* (SC6A9.34)	1	82 /768
Proteins of Unknown Function:			
putative protein[72]	*S. coelicolor* (SCD78.22)	1	33 /339
putative protein (mutT domain)[73]	*S. coelicolor* (SCI35.14)	1	167 /180
putative protein[73]	*S. coelicolor* (SCI35.26)	1	44 /175
putative protein[74]	*S. coelicolor* (SC1C3.22)	1	56 /442
putative integral membrane protein[75]	*S. coelicolor* (SC3C8.03)	1	249 /589
putative protein[75]	*S. coelicolor* (SC3C8.05)	1	7 /127
putative protein[75]	*S. coelicolor* (SC3C8.06)	1	45 /188
putative protein[75]	*S. coelicolor* (SC3C8.21)	1	44 /175
putative protein[76]	*S. coelicolor* (SC4G2.12)	1	21 /605
putative secreted protein[77]	*S. coelicolor* (SC8A6.16)	1	17 /380
putative protein[78]	*S. coelicolor* (SC9C7.12)	1	12 /197
96 targets		**106**	**1-25 (31)** **26-50 (16)**

Reference: Author and Genbank Accession Number

[1] Gandecha and Cundliffe (1996), #X97721
[2] Distler *et al.* (1992),
Vogtli and Hutter (1987), # AJ006985
[3] Zalacain *et al.* (1986), #X03615
[4] Demidchuk et al. (1997), #Y15838
[5] August *et al.* (1994), #L29247
[6] Hara and Hutchinson (1990)
[7] Epp *et al.* (1987), #M16503
[8] Fernandez-Moreno *et al.* (1991), #M64683
[9] Martin (1998), #AJ001743
[10] Beyer *et al.* (1996), #AJ006985
[11] Geistlich *et al.* (1992), #X63451
[12] Scheu *et al.* (1997), #X91393
[13] Lombo *et al.* (1999), #AF056309
[14] White and Bibb (1997), #AI021409
[15] Perez-Llarena *et al.* (1997), #AF073897
[16] Lyutzkanova *et al.* (1997), #U70376
[17] Distler *et al.* (1987), #Y00459
[18] Raibaud *et al.* (1991), #D00609
[19] Ruan *et al.* (1997), #AF007101
[20] Otten *et al.* (1995), #L37338
[21] Ylihonka *et al.* (1996b), #AJ224512
[22] Babcock and Kendrick (1990),
McCue et al. (1996), # M32687
[23] McCue *et al.* (1992), McCue *et al.* (1996),
Redenbach et al. (1996), #M80614
[24] Ueda *et al.* (1993), #AB006206
[25] Redenbach *et al.* (1996), #AL031350
[26] Servin-Gonzalez *et al.* (1997), #M86351
[27] Vomastek *et al.* (1998), #AJ223176
[28] Molnar *et al.* (1996), # X86780
[29] Tanaka *et al.* (1990), #D00671
[30] Arisawa *et al.* (1993), #D31821
[31] Swan *et al.* (1994), #L09654
[32] Tercero *et al.* (1996), #X92429
[33] Kakavas *et al.* (1997), #AF016585
[34] Watanabe *et al.* (1995), #D30815
[35] Gandecha *et al.* (1997), #X81885
[36] DeHoff *et al.* (1996), #U78289
[37] Distler *et al.* (1992), #AJ006985
[38] Decker *et al.* (1993), #M80674
[39] Sakata et al. (1992), #D11457
[40] Heinzel *et al.* (1988), #X53527
[41] Pissowotzki *et al.* (1991), #X62567
[42] Ruan, *et al.* (1997), #AF007101
[43] Schwecke *et al.* (1995),#X86780
[44] Coque *et al.* (1991), #X57310
[45] Peschke *et al.* (1995), #X79146
[46] Hara and Hutchinson (1992), #M93958
[47] Torkkell *et al.* (1997), #AJ224512
[48] Dickens *et al.* (1996), #U43704
[illegible]otamedi *et al.* (1997), #Y10438
[illegible]Arisawa *et al.* (1995), #D30759
[51] Brown *et al.* (1996), #U21728
[52] Schwartz *et al* . (1996), #X65195
[53] Homerova *et al.* (1992), #M82920
[54] Redenbach *et al.* (1996), #AL021529
[55] Redenbach *et al.* (1996), #AL022268
[56] Omer *et al.* (1990), #M322238
[57] Poehling *et al.* (1996), #X95915
[58] Garda *et al.* (1997), #U51222
[59] Kawamoto *et al.* (1993)
[60] Tsao *et al.* (1993), #X70255
[61] Wittmann *et al.* (1994), #U04629
[62] Robbins *et al.* (1992), #M82804
[63] Perito *et al.* (1994), #X73953
[64] Horii *et al.* (1990), #M31939
[65] Mori *et al.* (1997), #AB007189
[66] Katyama *et al.* (1995), #D50624
[67] Hagege *et al.* (1993), #AJ005260
[68] Hagege *et al.* (1994), #AJ005260
[69] Zotchev and Schrempf (1994), #X73563
[70] Bar-Nir *et al.* (1992), #M86368
[71] Redenbach *et al.* (1996), #AL031035
[72] Redenbach *et al.* (1996), #AL034355
[73] Redenbach *et al.* (1996), #AL031541
[74] Redenbach *et al.* (1996), #AL023702
[75] Redenbach *et al.* (1996), #AL023861
[76] Redenbach *et al.* (1996), #AL031371
[77] Redenbach *et al.* (1996), #AL031013
[78] Redenbach *et al.* (1996), #AL035161

almost 50% of the UUA codons are confined to the first 50 codons it seems likely that these codons together with the cognate *bldA*-tRNA function to modulate the expression of the target genes, similar to the manner that AGA/AGG codons and the *argU* tRNA modulate gene expression in *E. coli.*

In *S. coelicolor,* 20 *bldA* targets have been identified to date, and although the function of most of these genes has not yet been elucidated, the number of targets as well as the function of the characterized targets suggest that the *bldA* tRNA exerts its effect on multiple genes and pathways rather than at the level of a single global regulator. Three of the targets that have been characterized are implicated in antibiotic production. These are *act*II-ORF2, *act*II-ORF4, and *redZ*. As mentioned previously, *act*II-ORF4 and *redZ* are the pathway-specific regulators of the actinorhodin and undecylprodigiosin biosynthetic pathways respectively, and *act*II-ORF2 is involved in export of actinorhodin. While *bldA* mutants strains containing a cloned copy of *act*II-ORF4 on a low copy number vector produce only trace amounts of actinorhodin, strains containing a TTA→TTG mutant version of the *act*II-ORF4 gene produce wild-type levels of the antibiotic, although it is primarily cell-associated (Fernández-Moreno *et al.*, 1991). This provides evidence that the defect in actinorhodin production in a *bldA* mutant is due to inefficient translation of the single TTA codon in the ActII-ORF4 pathway-specific regulator, and the inability of the *bldA* strains containing the TTG corrected version of *act*II-ORF4 to export actinorhodin is consistent with a requirement for *bldA* for the translation of one of the genes implicated in actinorhodin export, *act*II-ORF2. Moreover, since *act*II-ORF4 is transcribed normally in *bldA* strains (White and Bibb, 1997) it provides further evidence that the effect of the *bldA* mutation is exerted at the level of translation of *act*II-ORF4.

Similarly, the dependence on *bldA* for undecylprodigiosin production is mediated by *redZ*, and multiple copies of *redZ* can restore undecylprodigiosin production to a *bldA* mutant (White and Bibb, 1997). However, *redZ* transcripts are more abundant in the *bldA* mutants as compared to wild-type, but since transcription of *redD*, which requires RedZ for activation, is severely reduced in *bldA* mutants it seems likely that RedZ negatively autoregulates its transcription and that the translation of *redZ* is defective in *bldA* mutants. Although none of the *S. coelicolor bldA* targets identified thus far are implicated in regulation or production of either methylenomycin or the calcium-dependent antibiotic, it seems likely that the pathway-specific regulators for these two antibiotic biosynthetic pathways will prove to be either directly or indirectly *bldA* dependent.

The only *S. coelicolor bldA* target that has been implicated in sporulation is *bldX*, which is the homologue of *S. griseus nrsA* gene (Babcock and Kendrick, 1990; McCue *et al.*, 1996). However, *nrsA* null mutations have only been successfully generated in *S. griseus* strains containing additional mutations that block sporulation at an early stage, suggesting that disruption of this gene is lethal in sporulation competent strains and this gene functions to prevent premature sporulation. If *bldX* functions in a similar manner in *S. coelicolor*, and it represents the only sporulation-specific *bldA* target, then the *bldA* mutation would likely be lethal as neither of the *bldX* nested open reading frames would be fully translated. Since *bldA* mutants are viable, there must be an additional *bldA* target involved in sporulation that negates the effect of *bldX*. A likely candidate for this additional *bldA* target is an extracellular signalling molecule since *bldA* mutants can be complemented extracellularly by *bld* mutants blocked at a later stage in development.

One of the genes encoding a response regulator *(amfR)* in *S. griseus,* and implicated in the regulation of sporulation, contains a TTA codon. While the *S. coelicolor* homologue, *ramR,* is not *bldA* dependent, there may be additional response regulators or sensor kinases in *S. coelicolor* that are *bldA* dependent. There are numerous additional *bldA* targets in *S. coelicolor* that have not yet been characterized, and while some of them, such as the putative extracellular hydrolase and putative extracellular protease, are likely involved in secondary metabolism, others, such as the putative ATP/GTP binding protein, putative integral membrane protein, putative secreted protein, or one of the numerous proteins of unknown function, may prove to be involved in sporulation or secondary metabolite production.

SCOPE OF THESIS

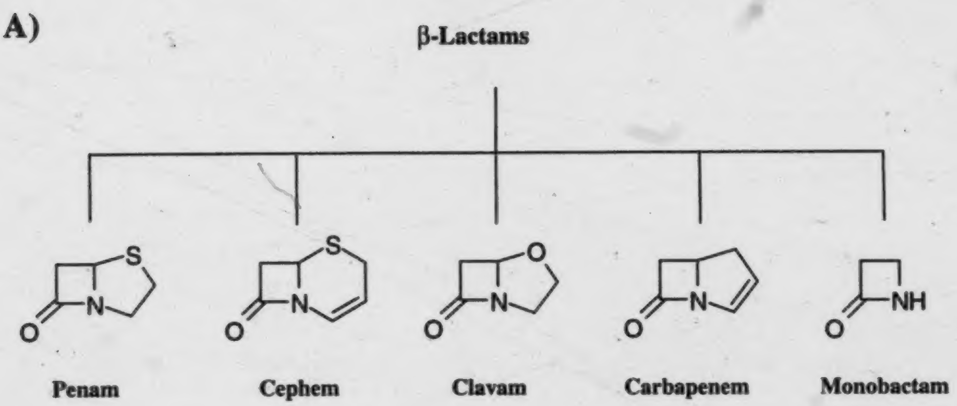
Since the *bldA* gene is required for sporulation and antibiotic production in *S. coelicolor,* the closely related *S. lividans*, and the more distantly related *S. griseus,* it seems likely that the *bldA* gene plays a similar role in all *Streptomyces* species. Furthermore, the distribution of TTA codons in genes that function relatively late in growth, and the bias of TTA codons near the start of genes, supports a role for *bldA* in the modulation of expression of genes that function late in growth. In order to provide support for the hypothesis that *bldA* plays a common role in the genus *Streptomyces,* it was necessary to identify the *bldA* gene in an additional species and to generate *bldA* mutant strains that are sporulation defective and antibiotic deficient. Since all of the *Streptomyces* species sporulate, the choice of strain in which to identify a *bldA* homologue was based on antibiotic class. *S. coelicolor* and *S. lividans* produce the anthracycline antibiotic actinorhodin which along with other antibiotics of the quinone moiety and macrolide antibiotics are collectively referred to as polyketide antibiotics, and constitute a major class of antibiotics in *Streptomyces* (Crandall and Hamill, 1986). *S. griseus* produces the aminocyclitol-amino-glycoside antibiotic streptomycin which is active against Gram negative organisms (Mansouri *et al.*, 1989) and belongs to the major structural class of streptomycete antibiotics referred to as aminocyclitols. This leaves the antibiotics of the major structural class ansamycins, β-lactams, nucleosides, peptides, polyenes, polyethers, and tetracyclines unrepresented with respect to dependence on *bldA* for expression so it seemed appropriate to identify the *bldA* gene in one of the species that produces one or more of these major classes.

β-lactams are among the most well-known antibiotics, consisting of penicillins, cephamycins, cephalosporins, carbapenems, clavams, and monobactams (Figure I.2) and are produced by fungi such as *Penicillium* and *Cephalosporium*, numerous *Streptomyces* species (Crandall and Hamill, 1986) as well as by numerous unicellular Gram-positive and Gram-negative bacteria (Jensen and Demain, 1993). β-lactam antibiotics have been used clinically for more than fifty years (Knowles, 1985), and are still widely prescribed. β-lactams target transpeptidases involved in peptidoglycan biosynthesis, and since peptidoglycan is a component of all bacterial cell walls but not of eukaryotic membranes, the antibiotic is both highly specific for prokaryotes (non-toxic) and has broad-spectrum activity. For these reasons β-lactams are still one of the preferred classes of antibiotics for clinical use, although increased bacterial resistance poses a significant problem. The most prevalent mode of resistance to β-lactams is by β-lactamases, enzymes that irreversibly inactivate β-lactam antibiotics. Initially, efforts to overcome resistance imparted by β-lactamases focused on identifying new compounds that were not substrates

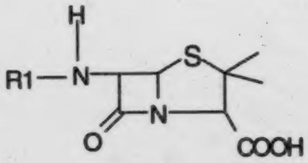
Figure I.2

Structure of β-lactam antibiotics. **(A)**. Types of β-lactams. **(B)** General structure of penicillins, cephamycins, and cephalosporins.

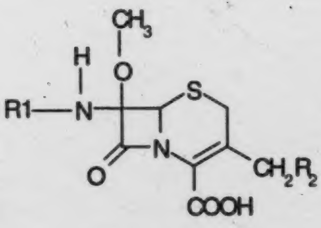
A)


β-Lactams
S
N
O
Penam
S
N
O
Cephem
O
N
O
Clavam
N
O
Carbapenem
NH
O
Monobactam

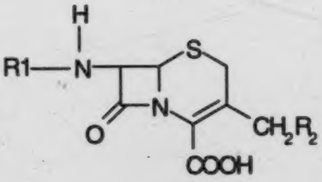
B)


H
R1
N
S
N
O
COOH

Penicillin


CH3
H
O
R1
N
S
N
O
CH2R2
COOH

Cephamycin


H
R1
N
S
N
O
CH2R2
COOH

Cephalosporin

for these enzymes, but as resistance to each new β-lactam has occurred and as most naturally occurring β-lactam antibiotics have been identified, the research has shifted to investigation of compounds that inhibit β-lactamases, and to investigation of the β-lactam biosynthetic pathways in order to maximize production and potentially produce novel compounds. Identification of the *bldA* gene in a β-lactam producing streptomycete may provide further insight into the regulation of antibiotic production and possible means of maximizing production, so a β-lactam producing strain seemed a logical choice for identification of a *bldA* homologue.

Although *S. griseus* strain NRRL 3851 produces the β-lactams cephamycin A and cephamycin B(Crandall and Hamill, 1986; Stapley *et al.*, 1972), the *bldA* gene was characterized in strain NRRL-B-2682 which doesn't produce any of the β-lactam compounds. Rather than identify the *bldA* homologue in *S. griseus* NRRL 3581, which is closely related to strain NRRL-B-2682, it was decided to identify the *bldA* homologue in a more distantly related strain. The strains chosen for further study were *S. lipmanii,* which produces the β-lactam 7-methoxycephalosporin C and clavaminic acid (Crandall and Hamill, 1986) (Alexander and Jensen, 1998), *S. jumonjinensis,* which produces cephamycin C and clavulanic acid (Ward and Hodgson, 1993), and *S. clavuligerus,* which produces cephamycin C, clavulanic acid, and numerous clavam compounds (Brown and Evans, 1979; Crandall and Hamill, 1986; Pruess and Kellett, 1983; Reading and Cole, 1977). Of these three species, the antibiotic biosynthetic pathways have been investigated primarily in *S. clavuligerus,* and this species was eventually chosen for further study, so the antibiotics and biosynthetic pathways of the other two species will not be discussed further.

Cephamycin C is structurally related to other cephamycins as well as to penicillins and cephalosporins, and all are formed from the initial condensation of L-α-aminoadipic acid, L-cysteine, and L-valine (Aharonowitz *et al.*, 1992; Jensen and Demain, 1993). In fungi, this is the first step in the production of penicillins and cephalosporins as the precursor L-α-aminoadipic acid is produced as an intermediate in lysine biosynthesis. In *Streptomyces* and other prokaryotes, lysine has to be converted to L-α–aminoadipic acid in a two-step process (Romero *et al.*, 1997). The first step is carried out by lysine-ε-aminotransferase (LAT), and is considered the first enzyme in the β-lactam biosynthetic pathway (Madduri *et al.*, 1991; Tobin *et al.*, 1991). Consistent with this, the *lat* gene maps to the β-lactam biosynthetic gene clusters in β-lactam producing *Streptomyces* species. The second step in the conversion of lysine to L-α-aminoadipic acid is catalyzed by piperideine-6-carboxylate dehydrogenase (PCDH), encoded by the *pcd* gene (de la Fuente *et al.*, 1997; Martin, 1998). The condensation of L-α-aminoadipic acid, L-cysteine and L-valine is carried out by ACV synthetase (ACVS), the product of the *pcbAB* gene, and is followed by cyclization of the linear tripeptide to isopenicillin N by the isopenicillin N synthase (IPNS), *pcbC* gene product (Aharonowitz *et al.*, 1992). The genes *lat, pcbAB,* and *pcbC* in prokaryotes and *pcbAB* and *pcbC* in fungi represent the early genes in the biosynthesis of β-lactams, as after this point the penicillin biosynthetic pathway branches off from the cephalosporin and cephamycin pathways. Most of the remaining genes in the cephamycin C biosynthetic gene cluster have been identified and characterized (Figure I.3), and the reactions they catalyze are shown in Figure I.4. The biosynthetic cluster is regulated by the transcriptional activator CcaR, and the corresponding gene, *ccaR* also maps within the

Figure I.3

S. clavuligerus cephamycin C biosynthetic gene cluster. The location, orientation, and relative sizes of the genes are indicated by bold arrows. Transcripts are indicated below the respective genes. The dashed lines indicate transcripts with incompletely characterized 5' or 3' ends.

Streptomyces clavuligerus **Cephamycin C Biosynthetic Gene Cluster**

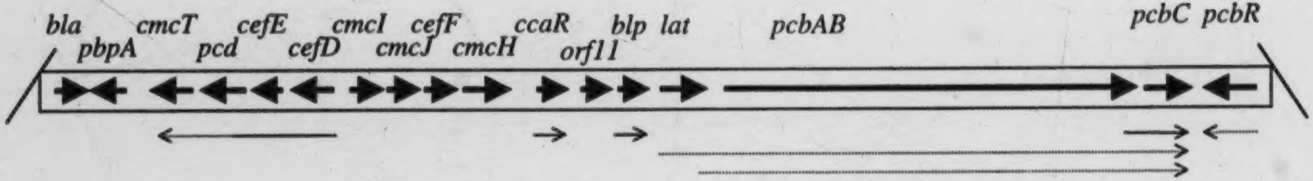
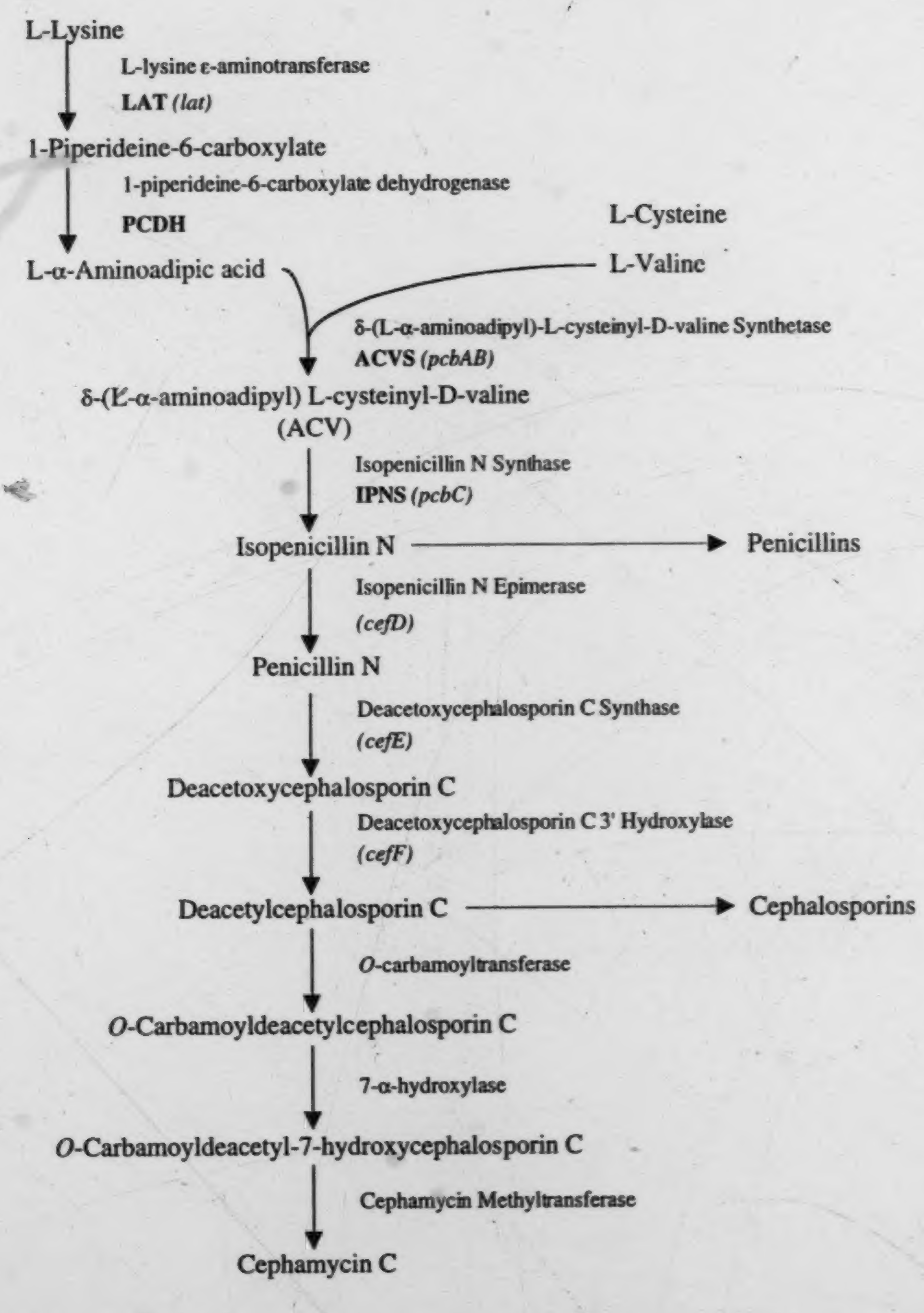

bla
pbpA
cmcT
pcd
cefE
cefD
cmcI
cmcJ
cefF
cmcH
ccaR
orf11
blp
lat
pcbAB
pcbC
pcbR

Figure I.4

The biosynthetic pathway for cephamycin C in *S. clavuligerus.* The enzymes and the corresponding genes for each step are indicated. The branch-points for penicillins and cephalosporins are also indicated.

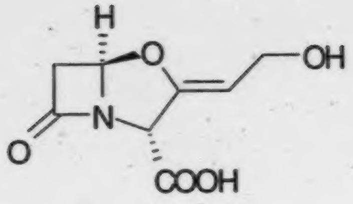

L-Lysine
L-lysine ε-aminotransferase
LAT (lat)
1-Piperideine-6-carboxylate
1-piperideine-6-carboxylate dehydrogenase
PCDH
L-Cysteine
L-α-Aminoadipic acid
L-Valine
δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine Synthetase
ACVS (pcbAB)
δ-(L-α-aminoadipyl) L-cysteinyl-D-valine
(ACV)
Isopenicillin N Synthase
IPNS (pcbC)
Isopenicillin N
Penicillins
Isopenicillin N Epimerase
(cefD)
Penicillin N
Deacetoxycephalosporin C Synthase
(cefE)
Deacetoxycephalosporin C
Deacetoxycephalosporin C 3' Hydroxylase
(cefF)
Deacetylcephalosporin C
Cephalosporins
O-carbamoyltransferase
O-Carbamoyldeacetylcephalosporin C
7-α-hydroxylase
O-Carbamoyldeacetyl-7-hydroxycephalosporin C
Cephamycin Methyltransferase
Cephamycin C

cluster (Perez-Llarena *et al.*, 1997). CcaR is a protein of 256 amino acids, and shows overall end-to-end homology to the products of the *act*II-ORF4 and *redD* genes, as well as to other transcriptional activators (Perez-Llarena *et al.*, 1997). CcaR is also a target for the action of *bldA*, as it contains a TTA codon in position 32. Although it has not yet been determined which promoters the CcaR protein activates, it seems likely that it regulates the expression of most of the genes, as LAT, ACVS, IPNS, and an enzyme involved in one of the middle steps, DAOCS (*cefE* gene product) are not produced in *ccaR* mutant strains (Alexander and Jensen, 1998). The transcriptional organization of the three early genes in cephamycin C biosynthesis has been characterized. *pcbC* is transcribed as a monocistronic transcript from a promoter located within the *pcbAB* gene as well as part of a longer polycistronic transcript originating from the *lat* promoter (Alexander, 1998; Petrich *et al.*, 1994; Petrich *et al.*, 1992). The *cefE* gene is cotranscribed along with *cefD* and possibly *pcd* and *cmcT* (Kovacevic *et al.*, 1990). The *cmcT* gene is another *bldA* target as codon 470 of the 523 amino acid protein is a TTA codon. The CmcT protein product shows homology to export and resistance proteins and is likely involved in export of one or more of the β-lactams. If as predicted, *ccaR* is translationally dependent on *bldA*, and the CcaR product is required for transcription of the cephamycin C biosynthetic gene cluster then none of the biosynthetic genes should be transcribed in a *bldA* mutant, and cephamycin C should not be produced.

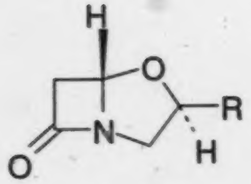
S. clavuligerus also produces clavulanic acid as well as several other clavam compounds (Figure I.5)(Brown and Evans, 1979; Reading and Cole, 1977). The clavams all have a bicyclic ring structure, but clavulanic acid differs in its stereochemical conformation (Janc *et al.*, 1993). Clavulanic acid is a weak antibiotic but is a potent

Figure I.5

Structure of clavulanic acid and clavam compounds.



Clavulanic acid



Clavam Compounds

β-lactamase inhibitor, as it irreversibly inactivates many β-lactamases (Reading and Cole, 1977). It is prescribed in combination with a β-lactam antibiotic, to combat infections caused by β-lactamase producing organisms (Foulstone and Reading, 1982; Hunter *et al.*, 1980; Rolinson, 1982). In contrast, the other clavam compounds have antifungal and/or antibacterial activity (Paradkar and Jensen, 1995; Pruess and Kellett, 1983). While little is known about the biosynthesis of the clavam compounds, their overall structural similarity to clavulanic acid suggests that their biosynthesis is similar to the early steps in clavulanic acid biosynthesis (Egan *et al.*, 1997; Janc *et al.*, 1993). Clavulanic acid biosynthesis is also poorly understood, although it has been determined that biosynthesis begins with the condensation of arginine and a three-carbon glycolytic intermediate (Valentine *et al.*, 1993). The biosynthetic gene cluster has been identified and maps adjacent to the cephamycin C gene cluster (Figure I.6) (Ward and Hodgson, 1993). Although the functions of most of the biosynthetic genes are currently unknown, a few of the biosynthetic genes have been characterized. *bls* encodes β-lactam synthetase which catalyzes the conversion of N^2-(2-carboxyethyl)arginine to deoxyguanidinoproclavaminate (Bachmann *et al.*, 1998). *cla* or *pah* encodes proclavaminate amidinohydrolase and catalyzes the conversion of guanidoproclavaminate to proclavaminic acid, and *cas2* encodes clavaminate synthase that converts proclavaminic acid to clavaminic acid (Figure I.7) (Aidoo *et al.*, 1994; Paradkar and Jensen, 1995; Wu *et al.*, 1995). Interestingly, both *cla* and *cas2* have homologues that map outside of the clavulanic acid cluster, as disruption of either gene results in loss of clavulanic acid production only under certain growth conditions. Since these growth condition correspond to conditions under which clavam compounds are not

Figure I.6

S. clavuligerus clavulanic acid biosynthetic gene cluster. The location, orientation, and relative sizes of the genes are indicated by bold arrows. Transcripts are indicated below the respective genes. The dashed lines indicate transcripts for which the 5' end has not been mapped whereas the solid lines indicate that the 5' end has been mapped.

Streptomyces clavuligerus Clavulanic Acid Biosynthetic Gene Cluster

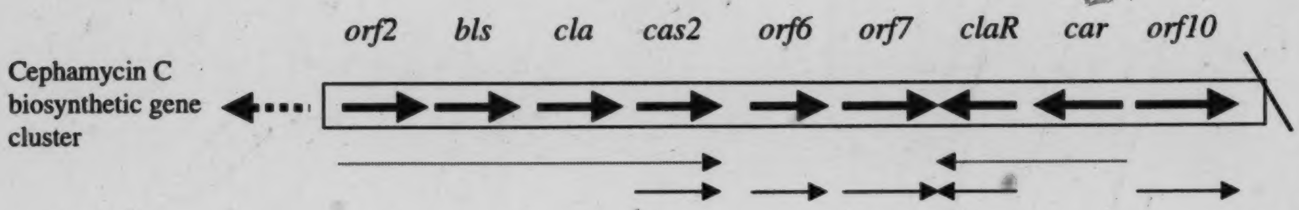
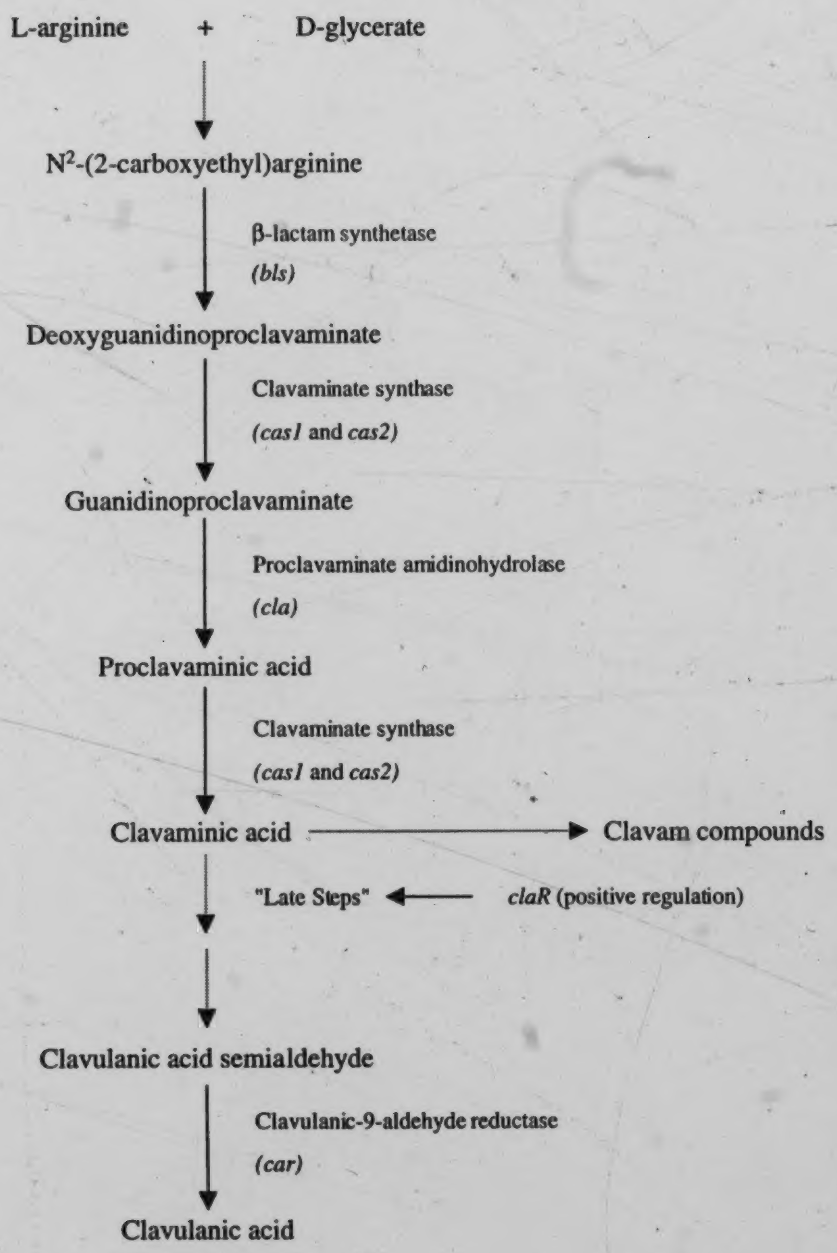

orf2
bls
cla
cas2
orf6
orf7
claR
car
orf10
Cephamycin C biosynthetic gene cluster

Figure I.7

The biosynthetic pathway for clavulanic acid in *S. clavuligerus.* The enzymes and the corresponding genes for each step that have been characterized are indicated.


L-arginine
+
D-glycerate
N²-(2-carboxyethyl)arginine
β-lactam synthetase
(bls)
Deoxyguanidinoproclavaminate
Clavaminate synthase
(cas1 and cas2)
Guanidinoproclavaminate
Proclavaminate amidinohydrolase
(cla)
Proclavaminic acid
Clavaminate synthase
(cas1 and cas2)
Clavaminic acid
Clavam compounds
"Late Steps"
claR (positive regulation)
Clavulanic acid semialdehyde
Clavulanic-9-aldehyde reductase
(car)
Clavulanic acid

produced (Paradkar and Jensen, 1995), it seems likely that the *cla* and *cas* homologues are involved in clavam compound biosynthesis. While the *cla* homologue has not yet been identified, the *cas2* homologue, *cas1*, has been sequenced (Marsh *et al.*, 1992).

The *bls, cla* and *cas2* gene products are involved in early steps in the biosynthesis of clavulanic acid, and the uncharacterized *orf-2* and *orf-6*, make up the remaining early genes (Hodgson *et al.*, 1995). The late genes are regulated by *claR* which encodes a positive activator (Paradkar *et al.*, 1998). The late biosynthetic steps likely represent the point where the biosynthetic pathways for clavulanic acid and other clavam compounds diverge, as disruption of *claR* prevents clavulanic acid biosynthesis under all conditions (Paradkar *et al.*, 1998). The only characterized late gene is *car*, which encodes clavulanic-9-aldehyde reductase and catalyzes the conversion of clavulanic acid semialdehyde to clavulanic acid (Perez-Redondo *et al.*, 1998). The biosynthetic cluster is regulated by *ccaR* (Perez-Llarena *et al.*, 1997), indicating that clavulanic acid production should also be regulated by *bldA*.

ccaR homologues have been identified by Southern hybridization in *S. cattleya, S. griseus, S. jumonjinensis,* and *S. lipmanii* (Alexander and Jensen, 1998), suggesting that β-lactams are regulated in a number of species by a *ccaR* homologue. In addition, cephamycin and clavulanic acid 'super-clusters' have been identified in *S. jumonjinensis* and *S. katsurahamanus* as well as in *S. clavuligerus* (Ward and Hodgson, 1993), and the clustering of these two pathways, along with the presence of a *ccaR* homologue in *S. jumonjinensis*, indicate that these two pathways are likely co-ordinately regulated by a *ccaR* homologue. This raises the possibility that β-lactam antibiotic production in *Streptomyces* is *bldA* dependent, through the translation of the pathway-specific activator, CcaR. Although clavam compound production has not been tested in a *ccaR* mutant, it seems unlikely that their biosynthesis is regulated by *ccaR,* as clavam compounds are not produced on starch asparagine medium, whereas CcaR and clavulanic acid are produced under these same conditions. Also the physical separation of the cephamycin – clavulanic acid super-cluster from the clavam biosynthetic genes indicates that the pathway-specific CcaR is unlikely to regulate transcription of the distant clavam cluster. If production of these compounds is *bldA* dependent, it is likely mediated through the translation of an as yet unidentified regulator.

The last two antibiotics produced by *S. clavuligerus* are holomycin and tunicamycin (Kenig and Reading, 1979; Okamura *et al.*, 1977). Holomycin is active against Gram-positive bacteria and tunicamycin is active against Gram-positive bacteria, fungi and viruses (Takatsuki and Tamura, 1971). At present, nothing is known about the biosynthesis or regulation of these two compounds, so their production will not be the focus of this study.

II. MATERIALS AND METHODS

II.1 MATERIALS

Restriction endonucleases were obtained from either Boehringer Mannheim, Laval, Quebec or New England Biolabs Ltd., Mississauga, Ontario. Polynucleotide kinase (from T4 infected *Escherichia coli* B cells), T4 DNA ligase, Klenow fragment of *E. coli* DNA polymerase, RNase-free DNase, *S1* nuclease and Avian Myeloblastosis Virus (AMV) reverse transcriptase were obtained from Boehringer Mannheim. Thermostable DNA polymerases Taq and Expand High Fidelity (a mixture of Taq and Pwo polymerase) were obtained from Boehringer Mannheim. Deep Vent thermostable polymerase was obtained from New England Biolabs. Amplitaq was purchased from Perkin Elmer Cetus, Norwalk, Connecticut, USA. Sequenase, a chemically modified form of T7 DNA polymerase, as well as dideoxyribonucleoside triphosphates and the modified nucleoside triphosphate 2'-deoxyribo-7-deazaguanosine-5'triphosphate, were purchased from United States Biochemicals, Cleveland, Ohio, USA. Deoxyribonucleoside triphosphates were purchased from Boehringer Mannheim. RNA guard was purchased from Pharmacia Biotech Inc., Baie d'Urfé, Quebec. Yeast tRNA was purchased from Sigma-Aldrich Canada, Ltd., Mississauga, Ontario. All enzymes were used according to specifications advised by the manufacturer.

Radioactively labelled [α-^{32}P]dATP and [γ-^{32}P]ATP were obtained from ICN Biochemicals, St. Laurent, Quebec. [α-^{32}P]dCTP, and [α-^{35}S]dATP were obtained from Amersham Canada, Oakville, Ontario. ThermoSequenase radiolabelled terminator cycle sequencing kit (^{33}P) was obtained from Amersham Life Sciences, Inc. Oakville, Ontario.

Oligonucleotide primers were obtained from the Department of Biological Sciences DNA Synthesis Laboratory, University of Alberta, Edmonton, Alberta, Canada. All of the oligonucleotide primers and their sequences are listed in Table II.1.1.

Apramycin was obtained from E. Seno, Eli Lilly and Company, Indianapolis, Ind., USA, Apralan (apramycin sulfate) was obtained from Provel, a division of Eli Lilly Canada Inc., Scarborough, Ontario. Thiostrepton was obtained from S. Lucania, Squibb and Sons, Inc., Institute for Medical Research, Princton, N.J., USA. Ampicillin, penicillin G, tetracycline, and kanamycin were obtained from Sigma-Aldrich Canada, Ltd. Clavulanic acid standards were obtained from Dr. S. E. Jensen, Department of Biological Sciences, University of Alberta, Edmonton, Alberta.

Amino acids were obtained from Sigma-Aldrich Canada, Ltd. Trypticase soy broth (TSB) was purchased from BBL, Becton Dickinson Microbiology Systems, Cockeysville, Maryland, USA. Yeast extract, Malt extract, ISP#3 (International *Streptomyces* Project Medium #3), Tryptone, Casamino acids, Bacto peptone and Bacto Agar were purchased from Difco Laboratories, Detroit, Michigan, USA. Mead Johnson Oatmeal Cereal (Pablum) and tomato paste were purchased locally at Safeway Canada.

Agarose was obtained from ICN Biochemicals, Inc., Aurora, Ohio, USA. Ultrapure agarose, ammonium persulfate (APS) and TEMED were obtained from Gibco BRL (Bethesda Research Laboratories), Gaithersbury, Maryland, USA. Premixed 40% acrylamide solutions (29:1, acrylamide:N,N'-methylene bisacrylamide; 19:1, acrylamide:N,N'-methylene bisacrylamide) were obtained from Bio-Rad Laboratories, Hercules, CA, and from Fisher Scientific, Fairlawn, New Jersey. Gene Clean was

Table II.1.1 Oligonucleotide Primer Sequences

Primer	Sequence 5'→3'	Region of homology	Use
BKL2	CATGGATCCACCCGGTAACTGATGCACC	upstream of *S. coelicolor bldA* gene	PCR of *bldA* gene for *bldA* probe
BKL15	GCCGCTGAGTCGCAACCA	downstream of *S. coelicolor bldA* gene	PCR of *bldA* gene for *bldA* probe
Universal	GTAAAACGACGGCCAGT	upstream of MCS (pUC and M13 based vectors)	sequencing, PCR, cloning
NTR1	TGCGTCAGCCCAGGTGG	upstream of *S. clavuligerus bldA* gene (579→563)	sequencing, primer extension *(bldA)*
NTR2	ACGGGGTTTCAGCCGCG	downstream of *S. clavuligerus bldA* gene (1227→1243)	sequencing
NTR3	AGAGCGGGACGCCGGGTAGTT	downstream of *S. clavuligerus bldA* gene (776→796)	sequencing of high G+C region
BKL84	CGGCCGTCATCCGCAGGT	upstream of ORFJI1 *(S. coelicolor)*	PCR and sequencing of ORFJI1
BKL85	CTTGTCCCGGCGCAGCTT	complementary to sequence of ORFJI1 (codons 92-97)	PCR and sequencing of ORFJI1, and ORF1
NTR21	ACACCGCAGCGAAGAGTGA	upstream of ORF1, complementary to bases 365→347	PCR and sequencing (ORF1), primer extension *(bldA)*
NTR5	GCGCGGTACCGCACTCTCCGTAACGAGA	3' end of *S. clavuligerus bldA* gene (*Kpn* I site at 5' end)	PCR of *bldA* downstream flanking sequence
NTR8	CGCCTCTAGACGATGAGGCGATCTTGAA	downstream of *S. clavuligerus bldA* gene (*Xba* I site at 5' end)	PCR of *bldA* gene for cloning
NTR9	CCGCGAATTCGCCATGGAACGCCTTGT	upstream of *S. clavuligerus bldA* gene (*Eco* RI site at 5' end)	PCR of *bldA* gene for cloning, S1 nuclease mapping, and promoter probe studies
NTR10	CGGAGCCGGACTCGAACC	complementary to the 3' end of the *bldA* tDNA	PCR of *bldA* gene probe, S1 nuclease mapping, and promoter probe studies
NTR20	GCTAGAATTCGGAGGCCGTCTCCGAGGA	upstream of *S. clavuligerus bldA* gene (*Eco* RI site at 5' end)	PCR of *bldA* gene for cloning, S1 nuclease mapping, and promoter probe studies
NTR6	GC(C/G)G(G/A) (G/A)GGGCGGAAT	homologous to 5' end of leucyl-tRNAs	PCR of leucyl-tRNAs
NTR7	(G/A) (G/A)GGC(C/G)GGACT(T/C)GA	complementary to 3' end of leucyl-tRNAs	PCR of leucyl-tRNAs
BKL42	CTCAAGCTAGCGCGTCTG	complementary to *leuU* α, from anticodon to D-loop	inverse PCR of leucyl-tRNA, southern and northern analysis
BKL43	GTGCTAGTGCCCTTTATC	homologous to *leuU* α, from anticodon loop to variable loop	inverse PCR of leucyl-tRNA
BKL48	CACTGCTGGTGGCTCCC	homologous to region upstream from *leuU* α	PCR of *leuU* α
BKL49	CTTCTCGTGCGACTCGG	complementary to region downstream from *leuU* α	PCR of *leuU* α
NTR11	GCCGAATTCACCTGGACCACCCACAAG	predicted to be homologous to region upstream from *leuU* β	PCR of *leuU* β
NTR12	GCCAAGCTTCACCGTATTTTCGCCCGT	complementary to region downstream from *leuU* β	PCR of *leuU* β
BKL5	TTAAGCTCGCCGTGTCT	complementary to *bldA* tRNA	Northern analysis of *bldA* transcripts
BKL53	CCCTGCAGTACCATCGGCGCT	complementary to *Streptomyces* 5S rRNA transcripts	Northern analysis (control for RNA loading)
Probe A	ACTATCGGAGAGGCCATG	complementary to +1-+17 of *cas2* transcripts	Northern analysis of *cas2* transcripts
BKL54	CCGCCTTCGCCACCGGT	complementary to *Streptomyces* 16 S rRNA transcripts	Northern analysis (control for RNA loading)
NTR16	GCGAGAATTCCACTCTTCGCTGCGGTGT	upstream of *S. clavuligerus bldA* gene (*Eco* RI site at 5' end)	PCR of probe for S1 nuclease mapping and promoter probe studies
NTR22	ACGTCTAGACCACCATCCGGGCCAGGA	complementary to 5' end of *S. clavuligerus bldA* tDNA (*Xba* I site at 5' end)	primer extension *(bldA)* and PCR of promoter probe insert
NTR23	CGCTCTAGAAGGCGTTCCATGGCGAA	upstream of *S. clavuligerus bldA* gene (552→534)	primer extension *(bldA)* and PCR of promoter probe insert
NTR24	GATCTGCAGCCAAGCTT	homologous to MCS of pIJ4083	PCR of insert in pIJ4083 (promoter probe studies)
NTR26	ATAGTTCATGTTGTCAGGT	complementary to MCS of pIJ4083	PCR of insert in pIJ4083 (promoter probe studies)

obtained from Bio 101 Inc., LaJolla, CA. Polyethyleneglycol (PEG) 1000 MW was obtained from NBS Biologicals, North Mymms, Hatfield, UK. Spermidine, lysozyme and dithiothreitol (DTT) were obtained from Sigma-Aldrich Canada, Ltd. Isopropyl β-D-thiogalactopyranoside (IPTG) and bovine serum albumin (Fraction V) were obtained from Boehringer Mannheim and X-gal was obtained from American BiOrganics, Niagara Falls, NY. DEAE membranes were purchased from Schleicher and Schuell, Inc., Keene, N.H. USA, and polyvinylidene difluoride (PVDF) Immobilon-P^{SQ} membranes were purchased from Millipore, Bedford, Mass. Western blot chemiluminenscence reagents were purchased from NEN Life Science Products, Boston, Mass. Rabbit α-CcaR primary antibody for Western blot analysis was generously provided by Dylan Alexander, University of Alberta, and the secondary antibody donkey, α-rabbit Ig horseradish peroxidase, was obtained from Amersham Canada, Oakville, Ontario. The *S. clavuligerus* strain containing a *ccaR::apr* disruption mutant was generously provided by Dr. S. E. Jensen, University of Alberta. Nuc-Trap® probe purification columns were obtained from Stratagene, La Jolla, CA. Micro Bio-Spin® 6 chromatography columns were obtained from Bio-Rad, Hercules, CA. Both purification columns were used according to the manufacturer's recommendations. Sephadex® G-50 was obtained from Pharmacia Biotech, Uppsala, Sweden.

All other chemicals used in this study were reagent grade.

The computer software programs used for the analysis of DNA and putative protein sequences were DNA Strider, which was designed and written by C. Marck (Commissariat a l'Energie Atomique, France) and FRAME, a program written by Bibb *et al.*, (1984) and adapted for the Apple MacIntosh by C. Jensen. Protein and nucleotide similarities were determined through the use of BLAST (Altschul *et al.*, 1990). Protein and nucleic acid sequence alignments, as well as transfer RNA sequence analysis and identification of inverted repeats, were performed with PC Gene, Intelligenetics Inc., written by Amos Bairoch, Department of Medical Biochemistry, University of Geneva, Switzerland, (1989), and Wisconsin Sequence Analysis Package, Version 8.1, Genetics Computer Group.

II.2 BACTERIAL STRAINS, PLASMIDS, PHAGES AND CULTURE CONDITIONS

II.2.1 Bacterial strains, plasmids and phages

The *Streptomyces* strains used are listed in Table II 2.1.1, the *Escherichia coli* and *Staphylococcus aureus* strains are listed in Table II 2.1.2 and the plasmids and phages used are listed in Table II 2.1.3.

II.2.2 Maintenance of *Streptomyces* strains

Streptomyces strains capable of sporulating were maintained as frozen glycerol stocks of spores. *bldA* strains were maintained as frozen mycelial stocks and since mycelial stocks were not viable indefinitely they were also maintained as lyophilized mycelial stocks. *Streptomyces coelicolor* and *Streptomyces lividans* were grown on R2YE agar (Hopwood *et al.*, 1985) at 30°C. *Streptomyces lipmanii* and *Streptomyces jumonjinensis* were grown on Tomato Oatmeal Agar (TOA) (2% tomato paste, 2% oatmeal pablum, 2.5% agar, pH 6.8) at 30°C. *Streptomyces clavuligerus* was grown at 28°C on sporulation medium (ISP#3) + additional Difco agar to a final concentration of 1.75 % agar for spore stocks and on Maltose Yeast Extract Malt Extract (MYM) agar (4% Maltose, 4% Yeast Extract, 10% Malt Extract, 2% agar) for frozen mycelial glycerol

Table II.2.1.1 *Streptomyces* Strains

Streptomyces strain	Genotype	Reference or Source
Streptomyces clavuligerus		
NRRL 3585	wildtype	Northern Regional Research Center, Peoria, Ill.
Δ*bldA* 4-1	Δ*bldA*, apramycin resistant	This work
ccaR::apr	*ccaR* disrupted with the *apr* gene	Alexander and Jensen, 1998
Streptomyces coelicolor **A3(2)**		
J1501	*hisA1, uraA1, strA1, pgl* , SCP1$^-$, SCP2$^-$	John Innes Institute; Chater, *et al.*, 1982
J1681	Δ *bldA, hisA1, uraA1, strA1, pgl,* SCP1$^-$, SCP2$^-$	John Innes Institute; Leskiw *et al.*, 1993
Streptomyces jumonjinensis		
NRRL 5741	wildtype	Northern Regional Research Center, Peoria, Ill.
Streptomyces lipmanii		
NRRL 3584	wildtype	Northern Regional Research Center, Peoria, Ill.
Streptomyces lividans **66**		
1326	SLP2, SLP3	John Innes Institute; Lomovskaya *et al.*, 1980
TK24	*str-6*, SLP2$^-$, SLP3$^-$	John Innes Institute; Hopwood *et al.*, 1983
TK64	*pro-2, str-6*, SLP2$^-$, SLP3$^-$	John Innes Institute; Hopwood *et al.*, 1983
J1725	*bldA39* in 1326-9 background	John Innes Institute; Leskiw *et al.*, 1991b

Table II.2.1.2 *Escherichia coli* and *Staphylococcus aureus* Strains

E. coli strain	Genotype	Reference or Source
DH5α	F$^-$, φ*80dlacZ* ΔM15, Δ(*lacZYA-argF*)*U169*, *deoR*, *recA1*, *endA1*, *hsdR17*(r_K^-, m_K^+), *phoA*, *supE44*, λ$^-$, *thi-1*, *gyrA96*, *relA1*	Life Technologies, Burlington, Ontario
DH5α F'	*F'*, φ*80dlacZ* ΔM15, Δ(*lacZYA-argF*)*U169*, *deoR*, *recA1*, *endA1*, *hsdR17*(r_K^-, m_K^+), *supE44*, λ$^-$, *thi-1*, *gyrA96*, *relA1*	Life Technologies, Burlington, Ontario
MV1193	Δ(*lac-proAB*), *thi*, *rspi* (str^r), *endA*, *sbcB15*, *hspR4*, Δ(*srl-recA*)306::*Tn10* (*tet*^{r}), *F'* [*traD36*, *proAB*, *lacI*^{q} *lacZ* ΔM15]	Zoller and Smith, 1987
XL1-Blue	*recA1*, *endA1*, *gryA96*, *thi1*, *hsdR17*(r_k^-, m_k^+), *supE44*, *relA1*, λ$^-$, *lac*$^-$, [F'*proAB*, *lacI*^{q} *lacZ* ΔM15, *Tn10* (*tet*^{r})]	Bullock *et al.*, 1987
ET12567	F$^-$*dam13::Tn9*, *dcm6*, *hsdM*, *hsdR*, *recF143*, *zjj202::Tn10*, *galK2*, *galT22*, *ara14*, *lacY1*, *xyl5*, *leuB6*, *thi1*, *tonA31*, *rpsL136*, *hisG4*, *tsx78*, *mtl1*, *glnV44*	gift from Doug MacNeil, Merck Sharp & Dohme Research Laboratories; MacNeil *et al.*, 1992
ESS	Cephamycin C sensitive indicator organism	A. L. Demain, Department of Biology, Massachusetts Institute of Technology, Boston, Mass.
***S.aureus* strain**		
N2	Penicillin resistant, clavulanic acid indicator organism	Department of Microbiology, University of Alberta Edmonton, Alberta

Table II.2.1.3 ***Escherichia coli*** **and** ***Streptomyces*** **plasmids**

Plasmids and phages	Relevant characteristic or genotype	Reference and Source
***Streptomyces* plasmids**		
pIJ486	high copy number cloning vector *(tsr)*	John Innes Institute; Ward *et al.*, 1986
pIJ4083	high copy number promoter probe vector *(tsr,* promoterless *xylE)*	John Innes Institute; Clayton and Bibb, 1990
SFGFP-pGRI:*ermE*	pIJ486 vector containing *gfp-mut3* gene under the control of the constitutive *ermE** promoter	Markus and Leskiw, 1997
***E. coli* plasmids**		
pUC118, pUC119	high copy number phagemid cloning vector *(Amp)*	J. Vieira, Waksman Institute of Microbiology, Rutgers University, Piscataway, N.J.; Vieira and Messing, 1987
pUC120	high copy number phagemid cloning vector *(Amp)* with *Nco* I cloning site	J. Vieira, Waksman Institute of Microbiology, Rutgers University, Piscataway, N.J.; Vieira and Messing, 1987
pUC120Ap(Nco)	pUC120 plasmid with apramycin resistance cassette with flanking *Nco* I restiction sites	gift from A. Paradkar, University of Alberta
pIJ4070	pUC18 derivative with *ermE** promoter mutant *(Amp)*	gift from M. Bibb, John Innes Institute; Bibb *et al.*, 1986
pIJ2925	pUC18 derivative with polylinker flanked by *Bgl* II sites	Janssen and Bibb, 1993
pWOR2925	pIJ2925 plasmid *(ccaR)* insert *(Nco* I site at 5' end of gene, TTA→CTA codon at L32)	gift from D. Alexander, University of Alberta
pAU5	pIJ2925 with *tsr* marker	Geibelhaus, *et al.*, 1996
pMDW	T7 bacteriophage promoter-based *E. coli* expression vector containing the *pcbC* gene	gift from M. Durairaj, University of Alberta
mut3gfp	pKEN expression vector containing green fluorescent protein (GFP) FACS-optimized mutant	B. Cormack, Department of Microbiology & Immunology, Stanford University School of Medicine, Stanford, CA.; Cormack *et al.*, 1996
***E. coli-Streptomyces* Shuttle Vectors**		
pSET152	Conjugal transfer from *E. coli* to *Streptomyces,* integrates into ΦC31att site *(Apr)*	Northern Regional Research Center, Peoria, Ill. Bierman *et al.*, 1992
pSET152(thio#2)	pSET152 with tsr marker cloned into the *Bam* HI site in the MCS	gift from L. Geibelhaus, University of Alberta
Bacteriophages		
M13mp18, M13mp19	M13 cloning and sequencing vector	Boehringer Mannheim; Yanisch-Perron *et al.*, 1985
M13KO7	helper phage	J. Vieira, Waksman Institute of Microbiology, Rutgers University, Piscataway, N.J.; Vieira and Messing, 1987

stocks and lyophilized mycelial stocks. Spore stocks were prepared by scraping spores from plates with a sterile spatula and resuspending in sterile milli-Q water. The spore suspension was placed in a water-bath sonicator to disperse the spores and filtered through sterile non-absorbent cotton wool packed into a 15 mL conical centrifuge tube with a hole drilled into the bottom. After filtration, the suspension was centrifuged at 2900 rpm for 10 minutes in a PR-J International Centrifuge (swinging bucket rotor), the pellet was resuspended in glycerol to give a final concentration of 20% (v/v), and the suspension was dispensed into either sterile 1.5 mL Eppendorf tubes or sterile 5 mL Bijou bottles. Working stocks were stored at $^{-}$20°C and permanent stocks were stored at $^{-}$70°C. For the mycelial stocks, the cultures were grown on cellophane discs placed on the surface of agar plates to prevent the substrate mycelia from penetrating the agar. Mycelia were scraped from the plates with a sterile spatula, suspended in sterile milli-Q water, homogenized, centrifuged at 2900 rpm for 10 minutes in a PR-J swinging bucket centrifuge, resuspended to a final concentration of 20% (v/v) glycerol, and stored at $^{-}$20°C or $^{-}$70°C as described above. Lyophilized mycelial stocks were prepared by scraping mycelia off the cellophane discs into 10% (w/v) Skim Milk. A small amount of the mixture was placed into the bottom of a freeze-dry vial, flash frozen in dry ice/ethanol and dried at $^{-}$70°C under vacuum for 24-48 hours. The vacuum-sealed vials were stored at $^{-}$20°C. Strains containing plasmids bearing thiostrepton or apramycin resistance determinants (tsr or apr^{R}) were propagated on media containing 5 µg/mL thiostrepton or 25 µg/mL apramycin for *S. clavuligerus* and 50 µg/mL thiostrepton or apramycin for *S. coelicolor* and *S. lividans* strains.

II.2.3 Maintenance of *E. coli* and *S. aureus* strains

All *E. coli* and *S. aureus* strains were grown at 37°C. *E. coli* MV1193 was maintained on minimal medium plates (1.21 mM $MgSO_4$, 0.014 M citric acid, 0.086 M dipotassium hydrogen orthophosphate, 0.025 M ammonium sodium hydrogen orthophosphate, 0.125% (w/v) glucose, 0.5 mg/L Vitamin B1 (thiamine), 1.5% a[illegible]r) at 4°C and XL1-Blue was maintained on LB agar plates (1% w/v tryptone, 0.5% (w/v) yeast extract, 1% (w/v) NaCl) containing tetracycline (12.5 µg/mL) at 4°C. *E. coli* strains containing plasmids and *S. aureus* N2 were maintained as frozen glycerol stocks. Sterile glycerol was added to 5 mL overnight LB or 2×YT (1.6% (w/v) tryptone, 1% (w/v) yeast extract, 0.5% (w/v) NaCl) cultures to a final concentration of 20% (v/v) glycerol. Strains containing plasmids were supplemented with 100 µg/mL ampicillin or 50 µg/mL Apralan. Cosmid containing strains were supplemented with 12.5 µg/mL tetracycli[illegible] The glycerol stocks were flash frozen in dry ice/ethanol and stored at ⁻70°C.

For the maintenance of M13 phages, aliquots of culture were transferred to Eppendorf tubes, centrifuged briefly to pellet the *E. coli* cells and stored at 4°C.

II.2.4 Culture conditions for isolation of DNA

For chromosomal DNA isolation, *Streptomyces* strains were grown according to Hopwood *et al.* (1985). Seed cultures were grown in universal vials containing springs (2-3 c[illegible] long with a diameter of 1 cm) and 10 mL of Trypticase Soy Broth (TSB): Yeast Extract-Malt Extract medium (YEME) (Hopwood *et al.*, 1985) at a ratio of 2:3 for 24-72 hours and used to inoculate 500 mL of YEME. For *S. coelicolor* and *S. lividans* , YEME broth was supplemented with Tiger's milk (L-arginine at 75 µg/mL, L-cystine, L-histidine, DL-homoserine, L-leucine, L-phenylalanine and L-proline at 56.25 µg/mL, adenine and uracil at 11.25 µg/mL and nicotinamide at 0.75 µg/mL)(Hopwood *et al.*, 1985), 0.5% glycine and 5 mM $MgCl_2$. For *S. clavuligerus, S. lipmanii,* and *S. jumonjinensis,* YEME contained 1% maltose instead of glucose, and was supplemented with 5 mM $MgCl_2$. All *Streptomyces* cultures were grown at 30°C and 250 rpm for 48 hours with the exception of *S. clavuligerus* which was grown at 28°C.

For the purpose of small-scale plasmid DNA isolation, *Streptomyces* strains were grown in 5-10 mL TSB supplemented with 1% (w/v) starch or maltose for *S. clavuligerus* and 1% (v/v) glycerol for *S. lividans.* Either thiostrepton (5 µg/mL for *S. clavuligerus* and 50 µg/mL for *S. lividans* and *S. coelicolor)* or Apralan (25 µg/mL for *S. clavuligerus* and 50 µg/mL for *S. lividans* and *S. coelicolor)* was added for plasmid maintenance. For large-scale plasmid preparations, cultures were grown as described for chromosomal preparations with the addition of antibiotic as described above. For small-scale plasmid isolations, plasmid-containing *E. coli* strains were grown overnight at 37°C on a tube roller in 5 mL of LB or 2×YT broth with 100 µg/mL ampicillin, 50 µg/mL Apralan, or 12.5 µg/mL tetracycline. For large-scale preparations, *E. coli* strains were grown overnight at 37°C and 250 rpm in 100 mL of antibiotic supplemented broth in 500 mL flasks.

For the purpose of single-stranded DNA isolation from *E. coli* cells containing phagemids, the strains were grown overnight on a tube roller at 37°C in 2×YT broth containing 100 µg/mL ampicillin. One hundred microlitres of overnight culture were inoculated into 9 mL of 2×YT broth containing 100 µg/mL of ampicillin, 500 µL of M13KO7 helper phage was added, and the culture was incubated at 37°C with vigorous shaking for 1.25 hours. Kanamycin was added to a final concentration of 70 µg/mL and the cultures were incubated overnight. For M13 phage isolation, single plaques were picked from plates into 4 mL of 2×YT broth with 40 µL of a stationary phase culture of DH5αF', MV1193, or XL1-Blue and incubated for 6 hours at 37°C.

II.2.5 Culture conditions for preparation of cell free extracts

For the preparation of cell free extracts, cultures were grown in TSB +1% (w/v) starch for *S. clavuligerus* and TSB + 1% (v/v) glycerol for *S. lividans*. Strains containing plasmids were supplemented with 5 µg/mL thiostrepton. Seed cultures were started from glycerol spore stocks or from plates and grown at 28°C *(S. clavuligerus)* or 30°C *(S. lividans)* for 48-72 hours at 250 rpm and used to inoculate 25 mL TSB starch or TSB glycerol broth. Samples were further incubated at 28°C or 30°C for 24-48 hours.

II.2.6 Culture conditions for preparation of protoplasts

For the purpose of generating protoplasts, *S. lividans* and *S. coelicolor* were grown according to Hopwood *et al* . (1985). Glycerol spore stocks were used to inoculate 10 mL seed cultures (TSB:YEME; described in II.2.4), which were grown at 30°C and 250 rpm for 48-72 hours. One to two millilitres of each seed culture was then used to inoculate 25 mL of YEME supplemented with 5 mM $MgCl_2$, 0.5% glycine and Tiger's milk (section II.2.4) in 250 mL flasks containing springs. Cultures were grown for 36-48 hours at 30°C and 250 rpm.

S. clavuligerus cultures were grown according to the procedure used by SmithKline Beecham Pharmaceuticals (personal communication) which is a modification of the procedure published by Bailey and Winstanley (1986). Seed cultures were grown initially in 10 mL TSB + 1% (w/v) maltose broth at 25°C (wildtype) and 28°C *(bldA)* for 48-72 hours. One millilitre of the seed culture was then used to inoculate 25 mL of TSB:YEME (10 mL of TSB, 15 mL YEME supplemented with 1% maltose instead of glucose and 5 mM $MgCl_2$) and incubated at 250 rpm for 18-24 hours at the temperatures indicated above.

II.2.7 Culture conditions for isolation of RNA

For the purpose of RNA isolation from surface grown cultures, *S. clavuligerus* was inoculated onto the surface of cellophane discs on Starch Asparagine agar [0.2% (w/v) L-asparagine, 2.1% (w/v) MOPS, 0.44% (w/v) KH_2PO_4, 1% (w/v) starch, 0.1% trace elements (trace element solution contains 1 mg/mL (w/v) each of $FeSO_4 \cdot 7H_2O$, $MnCl_2 \cdot 4H_2O$, $ZnSO_4 \cdot 7H_2O$ and 1.3 mg/mL (w/v) $CaCl_2 \cdot 3H_2O$); 2% (w/v) Agar, pH 6.8)]. The inoculum for wildtype cultures was a glycerol spore stock, and for the *bldA* strain the inoculum was a mycelial stock. The stocks were quantified so that each plate was inoculated with equivalent numbers of colony forming units (CFUs) sufficient to obtain a lawn of mycelia. The stocks were diluted in sterile milli-Q water prior to inoculation so that 100 µL of inoculum was spread on each plate. The plates were incubated at 28°C for 30-168 hours as required. For RNA isolation from liquid cultures, the inoculum was generated by growing both wildtype and *bldA* cultures on the surface of cellophane discs on MYM agar (0.4% (w/v) maltose, 0.4% (w/v) yeast extract, 1% (w/v) malt extract, 2% (w/v) Agar) at 28°C for 5-7 days. Since *S. clavuligerus* doesn't sporulate on MYM agar, both the wildtype and *bldA* strains grow similarly under these conditions. The mycelium was scraped off the cellophane discs into sterile milli-Q water, homogenized using a ground-glass homogenizer, and washed with 10.3% sucrose. The mycelia were resuspended in sterile milli-Q water and the OD_{600} determined. The mycelia were then used immediately to inoculate 100 mL volumes of TSB + 1% (w/v) starch in 500 mL flasks containing springs to an initial OD_{600} of 0.03. The cultures were then incubated at 28°C with shaking at 250 rpm for 24-72 hours as required.

For the isolation of RNA from *E. coli*, wildtype *E. coli* was inoculated into 5 mL LB or 2×YT broth and grown overnight at 37°C.

II.3 ISOLATION OF CHROMOSOMAL DNA, PLASMID DNA AND RNA

II.3.1 Isolation of chromosomal, plasmid, and single stranded DNA

Chromosomal DNA was isolated from *Streptomyces* strains according to one of the procedures described in Hopwood *et al.* (1985). Procedure 1 yielded the highest quality DNA and was used for the initial isolation of DNA from *S. clavuligerus, S. jumonjinensis, S. lipmanii,* and *S. coelicolor* . *S. lipmanii* was found to be extremely sensitive to the lysozyme treatment and was only incubated at 30°C in lysozyme buffer for 10-15 minutes, while the other strains were incubated at 30°C for 1 hour. Procedure 3, which can be performed much more rapidly than Procedure 1, yields DNA that is somewhat sheared although the average fragment size is reported to be greater than 40 kb (Hopwood *et al.*, 1985). This procedure was used to isolate chromosomal DNA for the purpose of verifying chromosomal gene replacements.

E. coli plasmid DNA and M13 RF DNA were isolated using a standard alkaline lysis protocol (Sambrook *et al.*, 1989) for small scale preparations. Cosmids were prepared by the same procedure except that Solution III was added immediately after the addition of Solution II. Large-scale isolations were also done using the alkaline lysis procedure, with further purification being achieved either by cesium chloride-ethidium bromide density gradients (Sambrook *et al.*, 1989) or Qiagen™ 500 tips. Qiagen™ 100 tips were used for small-scale plasmid isolation when high quality DNA was required.

Small-scale *Streptomyces* plasmid preparations were performed according to Sambrook *et al.* (1989) with the following modifications: 2 mg/mL lysozyme was added to Solution I and the cells were incubated at 37°C for 5-60 minutes and the volumes of Solution I, II, and III were doubled. For large-scale isolations, Qiagen™ 500 tips were used according to the manufacturer's recommendations with the following modifications; lysozyme was added to Buffer P1 to a final concentration of 2 mg/mL and incubated at 37°C for 30-60 minutes and the preparation was extracted with neutral phenol:chloroform (1:1) prior to loading on the column.

Single-stranded phagemid DNA and M13 phage DNA were isolated by the procedure of Messing and Vieira, (1982) with the following modifications: the phage were precipitated once from 1.3 mL of culture supernatant using 250 µL of 20% PEG 6000 in 3.5 M ammonium acetate at 4°C for 30-60 minutes; the pellets were dissolved in 100 µL of TE buffer (Hopwood *et al.*, 1985) and extracted twice with an equal volume of neutral phenol, followed by a single chloroform extraction; the DNA was precipitated with 0.5 volumes of 3.5 M ammonium acetate and 2 volumes of 95% ethanol; and the DNA was finally dissolved in sterile milli-Q water and the A_{260} determined.

II.3.2 Isolation of RNA

Streptomyces RNA was isolated by the modified procedure of Kirby *et al.* (1967) as outlined in Hopwood *et al.* (1985) with the following modifications: surface grown cultures were scraped off plates directly into modified Kirby mixture; the culture in modified Kirby mixture was vortexed for a total of 2 minutes using intervals of 30 seconds followed by 30 seconds incubation on ice; the mixture was transferred to 12 mL polystyrene tubes either with a 100-1000 μL micropipettor tip or a baked 5 mL glass pipette; centrifugation was done at 8,500 rpm in a Beckman J2-HS centrifuge with a JA-20 rotor; the phenol:chloroform extraction was repeated a minimum of two times (depending on the amount of interphase present); the first precipitation in 2-propanol was left for a minimum of 20 minutes on ice or stored at $^{-}70$°C; the mixture was centrifuged for 10 minutes to pellet RNA, washed with 1 mL of 95% ethanol and dissolved in 450 μL of diethyl pyrocarbonate (DEPC)-treated milli-Q water; the RNA was treated with 70 units of DNase at room temperature for 30 minutes, followed by two phenol:chloroform and chloroform extractions, and the RNA was precipitated in 2-propanol for a minimum of 20 minutes on ice. The purified RNA was dissolved in a minimum of 100 μL of DEPC-treated water and the A_{260} of a diluted sample was determined prior to storage as a 2-propanol precipitate at $^{-}70$°C.

RNA was extracted from *E. coli* by the procedure outlined by Frost *et al.* (1989). The procedure was modified slightly by the addition of a phenol:chloroform extraction following the phenol extraction, and the RNA was DNase treated as described above.

II.4 INTRODUCING DNA BY TRANSFORMATION

II.4.1 Transformation of *S. coelicolor* and *S. lividans*

Protoplasts were made and transformed as described in Hopwood *et al.* (1985). The formation of protoplasts was assessed visually by checking wet mounts of aliquots under a phase contrast microscope. Each 25 mL culture yielded 10 × 1 mL aliquots of protoplasts (approximately 4 x10^9 protoplasts/aliquot) which were stored at $^{-}70$°C as described. A single aliquot was used for each transformation and it was transferred to a 15 mL round-bottomed screw-capped tissue culture tube. If a frozen aliquot was used it was washed with 5 mL of P buffer to remove any nucleases released by freezing and thawing. An aliquot of 5-20 µL of a DNA solution in TE buffer was used for transformations although typically 10 µL (approximately 1 µg) was found to be sufficient. A solution of 25% PEG 1000 in P buffer was used to assist uptake of DNA. The PEG solution was mixed with the protoplasts by pipetting up and down three times, followed by the immediate addition of 5 mL of P Buffer. Following centrifugation at 3150 rpm (setting 5) in an International Clinical Centrifuge (swinging-bucket rotor), the protoplasts were resuspended in 600-1000 µL of P buffer, 100 µL aliquots were plated on 6-10 R2YE plates and the plates were incubated at 30°C. After 18 hours of incubation, the transformants were overlaid with 1 mL of a sterile suspension of 500 µg/mL thiostrepton in water to yield a final concentration of 25 mg of thiostrepton/mL of agar. The plates were further incubated at 30°C until individual transformants had reached sufficient size to subculture. Transformants were subcultured on R2YE agar containing 50 µg/mL of thiostrepton.

II.4.2 Transformation of *S. clavuligerus*

The procedure used to generate protoplasts and transform *S. clavuligerus* is based on the method of Hopwood *et al.* (1985) with the addition of heat treatment (Bailey and Winstanley, 1986) and additional modifications (SmithKline Beecham Pharmaceuticals, personal communication). The modifications to the procedure described by Hopwood *et* *al.* (1985) are as follows: mycelia were incubated in 1 mg/mL lysozyme in P buffer at 25°C instead of 30°C; an additional wash in 10 mL of P buffer was performed following filtration; the protoplasts were resuspended in 5-10 mL P buffer depending on the size of pellet and dispensed into 1 mL aliquots and stored at $^{-}70$°C. Protoplasts were transferred to 15 mL round-bottomed tissue culture tubes for transformation and, as described above, washed with P buffer if a frozen aliquot was used. Because of the highly active restriction-modification system of *S. clavuligerus*, covalently closed circular DNA for transformation was either isolated from a *dam*$^{-}$, *dcm*$^{-}$ *E. coli* strain (ET12567) or first passaged through *S. lividans*, a strain apparently devoid of restriction-modification systems. Additional procedures used to try to overcome the restriction system were a 10 minute heat shock at 45°C prior to the addition of plasmid DNA, and the addition of 10 µL of salmon sperm DNA (100 µg/mL) to the heat-shocked protoplasts immediately prior to the addition of plasmid DNA to provide a template for any remaining nucleases. A 10 µL aliquot (approximately 1 µg) of plasmid DNA was used for each transformation. After transformation, the protoplasts were resuspended in 500 µL of P buffer and 100 µL aliquots were plated onto five Modified R5B agar plates (10% sucrose; 11% L-glutamic acid; 0.1% casamino acids; 1% soluble starch; 0.025 M $MgCl_2$; 0.2 mM $MgSO_4$; 0.368% $CaCl_2$; 0.005% KH_2PO_4; 0.573% TES buffer, pH 7.2; 0.08 mg/L $ZnCl_2$; 0.4 µg/L $FeCl_3$; 0.02 µg/L $CuCl_2$; 0.02 µg/L $MnCl_2$; 0.02 µg/L $Na_2B_4O_7$; 0.02 µg/L $(NH_4)_6Mo_7O_{24}$; 2.75% agar). Transformants were incubated at 28°C for 48 hours and then overlaid with 1 mL of a suspension of 100 µg/mL thiostrepton or a solution of 500 µg/mL Apralan (or apramycin) in sterile water to yield a final concentration of 5 µg thiostrepton/mL of agar or 25 µg Apralan (or apramycin)/mL of agar. The plates were further incubated at 28°C until individual transformants had reached sufficient size to subculture. Transformants were subcultured on MYM agar containing 5 µg/mL thiostrepton or 25 µg/mL Apralan (or apramycin).

II.4.3 Transformation of *E. coli*

E. coli MV1193, XL1-Blue, DH5αF', and ET12567 competent cells were prepared according to the procedure of Tiong and Nash (personal communication). A 2.5 mL aliquot of an overnight broth culture of *E. coli* grown in 5 mL LB broth (supplemented with 12.5 µL/mL tetracycline as required for pilus maintenance) was used to inoculate 200 mL prewarmed YT broth. The culture was incubated at 37°C on a platform shaker until it reached an OD_{600} of 0.5 (approximately 135 minutes). The culture was separated into two 100 mL aliquots and centrifuged at 3000 rpm at 4°C in a Beckman J2-HS centrifuge and JA-10 rotor for 10 minutes. The pellets were resuspended in 40 mL of ice cold Solution A (30 mM CH_3COOK, 50 mM $MnCl_2$, 100 mM KCl, 15% glycerol, 10 mM $CaCl_2$), pooled and incubated on ice for 1 hour. The cells were centrifuged at 4°C for 10 minutes at 3000 rpm and the pellet was resuspended in 12 mL ice cold Solution B (10 mM MOPS (pH 7), 75 mM $CaCl_2$, 10 mM KCl, 15% glycerol). *E. coli* DH5α competent cells were obtained from GIBCO-BRL. All competent cells were aliquoted (200 µL or 50 µL for commercial cells) into individual 1.5 mL Eppendorf tubes and flash frozen in dry ice-ethanol and stored at -70°C. For transformations, competent cells were thawed on ice, plasmid DNA was added (0.5-15 µL volumes containing 50 ng-1.2 µg) and incubated for 30 minutes on ice. The cells were heat shocked for 45 seconds at 42°C (20 seconds at 37°C for commercial cells), incubated on ice for 2 minutes and LB broth added to 1 mL. The transformants were incubated at 37°C for 1 hour and then plated onto LB plates containing either 100 μg/mL ampicillin or 50 μg/mL Apralan. If blue-white selection was desired then the LB plates were also supplemented with Xgal and IPTG to a final concentration of 40 μg/mL Xgal and 0.1 mM IPTG. Transformants were incubated overnight at 37°C and subcultured on LB + antibiotic plates or into LB + antibiotic broth. M13 based vectors were transformed into competent *E. coli* as described for plasmid transformation. The transformed cells were added to 3 mL M13 soft agar (1% tryptone, 0.8% NaCl, 0.8% agar) at 45°C along with 10 μL 100 mM IPTG, 25 μL 40 mg/mL Xgal in dimethylformamide, and 200 μL of an overnight culture of *E. coli* MV1193, XL1-Blue or DH5αF' and poured over prewarmed M13 hard agar (1% tryptone, 0.8% NaCl, 1% agar) plates. After overnight incubation at 37°C, plaques were picked into 4 mL 2×YT containing 40 μL of an overnight culture of F'-containing *E. coli.*

II.4.4 Digestion, cloning and subcloning of DNA

Restriction enzyme digestions of plasmid and chromosomal DNA were carried out according to the recommendations of the supplier and Sambrook *et al.* (1989). Ligations were carried out with a 2:1 or 3:1 molar excess of insert : vector in 5-15 μL with 50-75 ng DNA/μL. The 5× ligation buffer consisted of 250 mM Tris-HCl, pH 7.6; 50 mM $MgCl_2$; 25% PEG 8000; 5 mM DTT. An aliquot of 100 mM ATP was added separately to a final concentration of 10 mM. Blunt-ended DNA fragments were ligated at room temperature (25°C) and cohesive ends were ligated at 15°C. A combination of blunt and cohesive end ligation was carried out overnight in a 4L beaker of water, which was chilled slowly from room temperature to 4°C. All ligations were incubated overnight.

II.5 TECHNIQUES USED FOR DNA ANALYSIS

II.5.1 Restriction fragment analysis by agarose gel electrophoresis

DNA fragments in the size range 0.7-8 kb were subjected to electrophoresis on 1% agarose gels using either a TAE buffer system (40 mM Tris-acetate, pH 8.0; 1 mM EDTA), or a TBE buffer system (90 mM Tris, 89 mM Boric acid, 2.5 mM Na_2EDTA). Appropriate molecular weight markers were either λ *Pst*I, λ *Eco*RI, or λ *Hin*dIII. One microlitre of loading dye (0.25% bromophenol blue, 40% w/v sucrose) was added to the samples to provide a visual assessment of migration. The DNA bands were visualised by straining in ethidium bromide and viewing on a UV transilluminator.

II.5.2 Restriction fragment analysis by polyacrylamide gel electrophoresis

DNA fragments in the size range 50-1000 bp were subjected to electrophoresis on 5% polyacrylamide gels (29:1, acrylamide:N,N'-methylene bisacrylamide) using a TBE buffer system. The molecular weight marker used was λ *Pst*I. Loading dye was added to the samples as previously indicated. The DNA bands were visualised on a UV transilluminator after staining in ethidium bromide.

II.5.3 DNA sequence analysis

DNA sequence analysis was performed using the chain termination method of Sanger *et al.* (1977) as modified by Tabor and Richardson (1987) for use with Sequenase™. Single stranded phagemid and phage template DNA was generated as described in section II.3.1. All reactions used α-^{35}S-dATP as the radioactive nucleotide. Sequencing ladders for primer extension were generated using a ThermoSequenase radiolabelled terminator cycle sequencing kit. The radioactive nucleotides used were α-^{33}P-ddNTPs.

Labelled fragments produced in the Sequenase™ sequencing reactions were separated by size on 6% denaturing polyacrylamide gels (19:1, acrylamide:N,N'-methylene bisacrylamide; 8.3 M urea) using the TBE buffer system (90 mM Tris, 89 mM Boric acid, 2.5 mM Na_2EDTA). ^{33}P-labelled fragments were separated on 6% denaturing polyacrylamide gels using a TTE buffer system (89 mM Tris, 29 mM Taurine, 0.54 mM Na_2EDTA). Electrophoresis was carried out at 35 Watts for 1.75-8 hours. Compressions in the DNA banding pattern were resolved by separating the sequencing reactions on sequencing gels containing 40% formamide or sending the templates to a sequencing service (Department of Biological Sciences sequencing service). Following electrophoresis, gels were soaked in fixative (10% methanol, 10% acetic acid) for 10 minutes, lifted onto 3MM Whatman No.1 filter paper and dried for 1-2 hours under vacuum at 80°C in a Bio-Rad Model 583 Gel Drier. Radioactive bands were visualised by exposing the sequencing gels to Kodak X-OMAT AR film at room temperature for 1-3 days followed by developing with a FUJI RGII X-ray film processor, or by exposing the gel to a phosphorimager screen and scanning in a Molecular Dynamics Model 445 SI phosphorimager.

II.5.4 Hybridization analysis

II.5.4.1 Preparation of Membranes for Colony Hybridization

Colony hybridization was used for preliminary screening of large numbers of *E. coli* clones, particularly when blue-white selection could not be used and when the desired event was expected to be rare. Colony hybridization was performed according to the Amersham manual (1985), with minor modifications. Circular Hybond-N nylon membranes were sectored with a pen and autoclaved. The membranes were placed on LB agar plates supplemented with the appropriate antibiotic. Colonies were picked onto the membrane (50-100/plate) as well as onto a duplicate master plate without the membrane using a sterile toothpick, and the plates were incubated overnight at 37°C. The membranes were removed from the plate and placed colony side up onto a piece of 3MM Whatman No.1 filter paper saturated with denaturing solution (1.5 M NaCl, 0.5M NaOH) for 5 minutes. The membranes were then transferred to a piece of filter paper saturated with neutralizing solution (1.5 M NaCl; 0.5 M Tris-HCl, pH 7.2; 0.001 M Na_2EDTA) for 5 minutes, then placed in a 2×SSC (0.3 M NaCl, 0.03 M trisodium citrate) solution and the cellular debris was rubbed off with a gloved hand. The filters were air-dried and UV cross-linked at 150 mJoules in a Bio-Rad GS Gene Linker.

II.5.4.2 Preparation of Dot Blots

This procedure was used to screen liquid cultures of cosmids. A 100 µL aliquot of each culture was spotted onto a damp Hybond-N nylon membrane placed in a dot blot apparatus. The membrane was left under vacuum in the apparatus for 5-10 minutes to prevent the samples from running. The membranes were then prepared as described for colony hybridization.

II.5.4.3 Preparation of membranes for Southern Hybridization

The procedure used to transfer DNA from agarose gels to nylon membranes is an adaptation of the method of Southern (1975) and is described by Hopwood *et al.* (1985).

DNA was separated on a 1% agarose gel and stained as described in section II.5.1. The gel was usually photographed with a ruler for later comparisons to the exposed film. The gel was then trimmed to the minimum size required, and soaked for 2 × 10 minutes in 0.25 M HCl with gently rocking (optional), 2 × 15 minutes in denaturing solution (1.5 M NaCl, 0.5M NaOH) with gentle rocking, rinsed three times in distilled water, and soaked in neutralizing solution (1.5 M NaCl; 0.5 M Tris-HCl, pH 7.2; 0.001 M Na_2EDTA) with gentle rocking for a minimum of 20 minutes. The gel was placed well-side down onto two pieces of 3MM Whatman No.1 filter paper saturated with 20×SSC (3 M NaCl, 0.3 M trisodium citrate). The gel and filter paper were placed onto a glass plate over a reservoir of 20×SSC, the bottom piece of filter paper was long enough that the ends rested in the 20×SSC and served as a wick. A Hybond-N nylon membrane was placed onto the gel, followed by two more pieces of Whatman filter paper, a stack of paper towels and a 0.5-1 kg weight. After overnight transfer the location of the wells were marked onto the membrane and the membrane was UV cross-linked at 150 mJoules on a Bio-Rad GS Gene Linker.

II.5.4.4 Generation of ^{32}P-labelled Probes

Probes generated from double-stranded DNA fragments were internally labelled with α^{32}P-dCTP or α^{32}P-dATP by the random primer labelling method described by Feinberg and Vogelstein (1983) and modified by Boehringer Mannheim. In a screw-capped tube, 9 µL (50 -500ng) of template DNA was heat denatured and chilled on ice. Two microlitres of hexanucleotide mix, 3 µL of 1.5 mM dNTP mix, 5 µL (50 µCi) of α^{32}P-dCTP or α^{32}P-dATP and 1 µL (2 units) of Klenow were added and incubated for 4 hours at 37°C or overnight at room temperature. Initially the radioactively labelled nucleotide used was α^{32}P-dATP and the dNTP mix consisted of dCTP, dGTP, and dTTP. This was later replaced with α^{32}P-dCTP to obtain a higher specific activity in the high G+C DNA of *Streptomyces* probes. The dNTP mix in these reactions consisted of dATP, dGTP, and dTTP. The labelled probes were separated from unincorporated nucleotides on a Sephadex G-50 column. One microlitre of purified labelled probe was counted in a Beckman LS 3801 scintillation counter to determine activity. Immediately prior to hybridization, the probes were heat denatured for 5-10 minutes at 95°C.

Oligonucleotide probes were 5' end-labelled with γ^{32}P-ATP by the protocol described by Chaconas and van de Sande (1980) and modified by Boehringer Mannheim. Twenty picomoles of oligonucleotide primer was incubated with 1 μL of 10× kinase buffer (0.5 M Tris-HCl, pH 8.0; 0.1 M $MgCl_2$; 50 mM DTT, 1 mM spermidine), 5 μL (50 μCi) of γ^{32}P-ATP in a total volume of 9 μL. One microlitre of 1/10 diluted polynucleotide kinase was added and the reaction was incubated at 37°C for 30 minutes. This was followed by the addition of another 1 μL aliquot of diluted kinase and incubation for a further 30 minutes. The dilution buffer for the polynucleotide kinase consisted of: 50 mM Tris-HCl, pH 8.2; 1 mM DTT; 0.1 mM EDTA, and 50% glycerol. Each probe was purified on a Sephadex® G-50, Nuc-Trap®, or Micro Bio-Spin® probe purification column and counted in a Beckman LS 3801 scintillation counter.

II.5.4.5 Filter Hybridization

Southern hybridization was carried out in heat sealable plastic bags (seal-a-meal) using 10-30 mL of hybridization solution, in Pyrex dishes with 50 mL of hybridization solution, or in glass hybridization tubes with 10-20 mL of hybridization solution. Colony hybridization was carried out in deep dish glass petri plates with 10 mL of hybridization solution. Prehybridization and hybridization solutions were made up of 3×SSC (0.45 M NaCl, 0.045 M trisodium citrate), 4× Denhardt's Solution (0.08% w/v Ficoll (MW 400,000), 0.08% w/v bovine serum albumin (Fraction V), 0.08% polyvinyl pyrrolidone (MW 360,000)), and 100 µg/mL of salmon sperm DNA. Prehybridizations were typically carried out overnight although this time was occasionally reduced to 6-8 hours and 4 hours for Southern and colony hybridizations, respectively. Hybridizations were carried out overnight with 2 million CPM probe/10 mL of hybridization solution. Hybridization temperatures were optimized for each probe. For probes 50 bp or longer the formula for calculating the T_m was $T_m = 81.5° + 16.6 \log M + 0.41 (\%G+C) - 500/n - 0.61 (\% \text{formamide})$, where M is the ionic strength (0.45 M for 3×SSC) and n is the length of the shortest duplex DNA segment (Hopwood *et al.*, 1985). The formula for calculation the T_d of oligonucleotide probes was $T_d = 4(G+C) + 2(A+T)$ (Hopwood *et al.*, 1985). Hybridizations were carried out 25° below T_m and 5° below T_d for duplexes without mismatches. If mismatches were expected the temperature was decreased by 1° for every 1% of mismatched bases for longer probes and 5° for every mismatched base for oligonucleotide probes. If the hybridization temperature was calculated to be greater than 65-70°C, formamide was added to the prehybridization and hybridization solutions to lower the T_m. Filters were washed at the hybridization temperature for 2 × 30 minutes in 2×SSC, 0.1% SDS and 2 × 30 minutes in 0.2×SSC, 0.1% SDS. The filters were then wrapped in Saran wrap and exposed to Kodak X-OMAT AT film at $^-70$°C and developed in a FUJI RGII X-ray film processor. Alternatively, the filters were exposed to a phosphorimager screen at room temperature and scanned with a Molecular Dynamics Model 445 SI phosphorimager. Blots were stripped by immersing in boiling 0.1% SDS and cooling to room temperature. This procedure was repeated until no signal could be detected upon exposure to film or a phosphorimager screen.

II.6 RECOVERY OF DNA FROM AGAROSE AND POLYACRYLAMIDE GELS

II.6.1 Recovery of DNA from agarose gels

Agarose gels were used to purify fragments greater than 1 kb in size. Purification from agarose gels was done using: Gene Clean, with ultrapure agarose and according to the manufacturer's recommendations; DEAE membranes (Schleicher & Schuell, Keene, NH), according to the manufacturer's recommendations; or by the trough method of Zhen and Swank (1993).

II.6.2 Recovery of DNA from polyacrylamide gels

DNA fragments up to 1 kb were isolated from polyacrylamide gels using the "Crush and Soak" method described in Sambrook *et al.* (1989), with the following modifications: the polyacrylamide slice was placed in an Eppendorf tube with 300-400 μL of elution buffer (0.5 M CH_3COONH_4; 1 mM EDTA, pH 8.0; +/- 10 mM $Mg(CH_3COO)_2$, and 0.1% SDS); the polyacrylamide was crushed with an Eppendorf tube grinder; and the DNA was eluted overnight at 37°C in a tube roller. To recover the supernatant, the Eppendorf tube was centrifuged at 15,000 rpm in an Eppendor centrifuge (Model 5415 C) for 5 minutes to pellet the acrylamide, and the supernatant was pipetted into a clean Eppendorf tube. The acrylamide was back-extracted with additional elution buffer, and the supernatant removed after vortexing and centrifugation. The pooled supernatants were centrifuged a second time to remove any remaining acrylamide, and the DNA was precipitated with ethanol.

II.7 POLYMERASE CHAIN REACTION (PCR) AND INVERSE PCR

II.7.1 Polymerase Chain Reaction

Polymerase chain reaction was used to amplify DNA fragments for cloning, and to generate probes for hybridization, sequencing, and transcription mapping. Standard reactions employed 100 µL volumes in 0.5 mL Eppendorf tubes and were carried out in a Techne PHC-2 thermocycler. Typical reactions contained 40 pmol of each primer; 0.2 mM dNTPs; 5% DMSO; 50 mM Tris-HCl, pH 9.2; 16 mM $(NH_4)_2SO_4$; 3.25 mM $MgCl_2$; 10 ng of plasmid template or 1 µg of chromosomal DNA template, and 0.7 µL (2.5 units) of Expand Hi-fidelity polymerase. Reactions were overlaid with 2 drops of mineral oil and denatured for 5 minutes at 95°, followed by 30 cycles of 30 seconds at 95°C, 30 seconds at 45-65°C, and 1 minute at 72°C. The annealing temperature was adjusted to 5°C below the T_d of the primer with the lower T_d. Some amplification reactions used Taq polymerase with a buffer containing 10 mM Tris-HCl (pH 8.3), [illegible].5 mM $MgCl_2$, 50 mM KCl, 0.1 mg/mL gelatine, or deep vent polymerase and a buffer containing 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 20 mM Tris-HCl (pH 8.8), 2-3 mM $MgSO_4$, 0.1% Triton X-100.

II.7.2 Inverse Polymerase Chain Reaction

Inverse polymerase chain reaction was used to amplify sequences flanking the *S. coelicolor leuU* tRNA. The template for inverse PCR was prepared by treating 1 µg of *SstII*-digested *S. coelicolor* J1681 DNA at 25°C with 0.02 U/mL T4 DNA ligase. PCR was performed with 40 pmol each of the primers BKL42 (5'-CTCAAGCTAGCGCGTCTG-3') and BKL43 (5'-GTGCTAGTGCCCTTTATC-3'), under conditions described above with

2% DMSO instead of 5% DMSO. The reaction mixture was denatured for 5 minutes at 95°C followed by 30 cycles of 95°C for 30 seconds, 52°C for 30 seconds, and 72°C for 1 minute.

II.8 TECHNIQUES FOR RNA ANALYSIS

II.8.1 Denaturing Polyacrylamide Gel Electrophoresis

RNA transcripts in the size range 50-800 bp were subjected to electrophoresis on 8% denaturing polyacrylamide gels (29:1, acrylamide:N,N'-methylene bisacrylamide; 8 M urea) using a TBE buffer system. RNA samples (5-10 µg) in 2-propanol were precipitated, washed with 80% ethanol and dissolved in 4 µL of loading dye (98% deionized formamide; 10 mM EDTA, pH 8.0; 0,025% xylene cyanol; 0.025% bromophenol blue). The molecular weight markers used were *Hpa*II-digested pBR322 and MW Marker V (Boehringer Mannheim; pBR322 DNA cleaved with *Hae*III) in loading dye. Samples and molecular weight markers were denatured for 5-10 minutes at 95°C, then chilled on ice. An aliquot containing 300-500 ng of molecular weight marker was loaded on each gel. The gels were electrophoresed at 120-150 volts for 4-6 hours or at 75 volts for 12 hours, stained in ethidium bromide in TBE buffer and photographed on a UV transilluminator.

II.8.2 RNA Agarose Gel Electrophoresis

RNA transcripts greater than 800 bases were subjected to electrophoresis on 1.25% Agarose gel using a 10 mM sodium phosphate buffer system. Forty microgram samples of RNA in 2-propanol were precipitated, washed with 80% ethanol and dissolved in 2.5 µL of DEPC-treated water. The samples were denatured with 2 µL glyoxal, 6 μL DMSO and 1.5 μL 80 mM sodium phosphate buffer, pH 6.5 for 60 minutes at 50°C, and chilled on ice (Williams and Mason, 1985). The molecular weight markers used were Molecular Weight Marker III (Boehringer Mannheim; λ-DNA, cleaved with *Hind*III), λ-*Bst*EII , and λ *Pst*I . Aliquots of 250-500 ng of molecular weight marker were denatured with deionized glyoxal and DMSO as described for the RNA samples. Three microlitres of loading dye (50% gylcerol; 10 mM $NaPO_4$, pH 7; 0.4% bromophenol blue) was added to RNA samples and molecular weight markers prior to loading. Denatured samples were separated by electrophoresis at 4 V/cm for 4 hours with recirculation.

II.8.3 RNA electroblotting onto nylon membranes from polyacrylamide gels

Prior to transfer, polyacrylamide gels were soaked for 3 × 10 minutes in 0.5% TBE to remove the urea. Transfer to Hybond-N nylon membrane was done using the TransBlot SD Semi-dry Transfer Cell (Bio-Rad) for 30-45 minutes at 3 mA/cm^2, under conditions recommend by the supplier. After transfer, the wells were marked on the membrane and the membrane was UV cross-linked at 150 mJoules in a Bio-Rad GS Gene Linker.

II.8.4 Northern transfer of RNA from Agarose gels

RNA separated on agarose gels was transferred to Hybond-N nylon membranes by capillary blotting overnight in 20×SSC. After transfer, the wells were marked on the membrane, the membrane was UV cross linked at 150 mJoules in a Bio-Rad GS Gene Linker and baked in an 80°C vacuum oven for 1-2 hours. The molecular weight marker lanes were removed and stained with 0.2% methylene blue; 0.2 M CH_3COONa, pH 4.7, (Miller, 1987) and washed with water.

II.8.5 Generation of ^{32}P-labelled probes

Probes generated from double-stranded DNA fragments were internally labelled with α^{32}P-dCTP or α^{32}P-dATP by the random primer labelling method described in section II.5.4.4.

Oligonucleotide probes and double-stranded DNA probes for S1 nuclease mapping were 5' end-labelled with γ^{32}P-ATP by the protocol described in section II.5.4.4.

II.8.6 Northern Analysis

Northern hybridization was carried out in glass hybridization tubes as described in section II.5.4.5. Ten to twenty millilitres of hybridization solution were used and the membranes were incubated in a hybridization oven. Prehybridizations were carried out overnight. Hybridizations were carried out overnight with 2 million CPM probe /10 mL of hybridization solution. The hybridization was optimized for each probe as previously described. For hybridizations with longer probes, formamide buffer (50% formamide, 6×SSC, 1×Denhardt's, 0.1% SDS) (Perez-Llarena *et al.*, 1997) was used to lower the T_m. Filters were washed at the hybridization temperature for 2 × 30 minutes in 2×SSC, 0.1% SDS and 2 × 30 minutes in 0.2×SSC, 0.1% SDS. The filters were wrapped in Saran wrap and exposed to Kodak X-OMAT AT film at $^{-}70$°C and developed in a FUJI RGII X-ray film processor, or exposed to a phosphorimager screen at room temperature and the images scanned in a Molecular Dynamics Model 445 SI phosphorimager. Blots were stripped by washing at 65°C in 0.5 mM Tris-HCl, pH 8.0; 2 mM EDTA, and 0.1×Denhardt's as recommended in the Amersham Manual. This was repeated until no signal could be detected by exposure to a phosphorimager screen overnight.

II.8.7 *S1* nuclease protection assays

S1 nuclease protection assays were performed to identify the transcription start point of the *bldA* transcript. The probes used for *S1* mapping were generated by PCR using primers designed to span the putative transcription start site. The primer homologous to the region upstream of the *bldA* tDNA (see section III.11) had a non-homologous 5' end, and was used to generate a probe with one non-homologous end. The purpose of this was two-fold: it allowed discrimination between probe-probe reannealed fragments and full-length protected fragments as the full-length protected fragments would be slightly shorter than the probe-probe reannealed fragments due to cleavage of this non-homologous extension on the probe by *S1* nuclease; it also functioned to prevent interference by antisense or other transcripts running divergent to the *bldA* tDNA gene as the labelled end of this probe strand would be removed by *S1* nuclease. RNA samples (10-40 µg) were hybridized to 50,000-200,000 CPM of ^{32}P end-labelled probe in formamide hybridization buffer (3.2 mM PIPES buffer, pH 6.4; 0.4 M NaCl; 1 mM EDTA; and 80% v/v deionized formamide). The samples were incubated at 85°C for 30 minutes, then cooled slowly to the annealing temperature (5° above the T_m of the length of the expected DNA/RNA hybrid). The samples were chilled on ice and treated with 1 unit of *S1* nuclease in digestion buffer (0.28 M NaCl; 30 mM CH_3COONa, pH 4.4; 4.5 mM $(CH_3CO_2)_2Zn$, and 20 µg partially-cleaved denatured calf thymus DNA) for 45 minutes at 37°C. A solution of 2.5 M CH_3COONH_4 and 0.05 M EDTA was added to terminate the reactions. The samples were extracted with phenol:chloroform, and then chloroform and precipitated in 2-propanol. The pellets were dissolved in 3 µL of loading dye (98% deionized formamide; 10 mM EDTA, pH 8.0; 0.025% xylene cyanol; and 0.025% bromophenol blue) and denatured for 5-10 minutes at 95°C prior to loading on a 6% denaturing polyacrylamide gel. Sequencing ladders for the identification of each transcript start point were performed using the PCR primer that was predicted to hybridize within the transcript. The single-stranded phagemid templates for the sequencing ladders were generated as described in section II.3.1.

II.8.8 Primer Extension

Primer extension reactions were performed according to the procedure outlined by Penfold *et al.* (1996) to identify the *S. clavuligerus bldA* transcription start sites. Primers were designed to hybridize approximately 100 bases from the proposed 5' end of the transcript. Fifty picomoles of primer was end-labelled with γ-^{32}P-ATP as previously described (Section II.5.4.4). Approximately 6×10^5 cpm of primer was used to anneal to 30 µg of RNA in 1 M NaCl; 167 mM Hepes, pH 7.5; and 33 mM EDTA, pH 8.0. The RNA and primer were incubated at 80°C for 5 minutes, then at 37 °C for 1 hour. The reactions were ethanol precipitated with 95 % ethanol, washed with 95 % ethanol and air dried for 20 minutes. The primer-annealed RNA was dissolved in 25 µL of reverse transcriptase mix (55 µM dNTPs; 50 mM Tris-HCl, pH 8.0; 5 mM $MgCl_2$; 5 mM DTT; 50 mM KCl; 50 µg/mL bovine serum albumin (Fraction V), and 0.5 µL of RNA guard). One microlitre (25 units) of AMV reverse transcriptase was added and the reaction was incubated at 42°C for 1 hour. One microlitre of 1 mg/mL RNase A was added and the reaction incubated at 37°C for 10 minutes, and then ethanol precipitated. The primer extension products were dissolved in 10 µL of loading dye (98% deionized formamide; 10 mM EDTA, pH 8.0; 0,025% xylene cyanol; and 0.025% bromophenol blue), heat denatured and 2 μL were loaded on a TTE-buffered 6% denaturing polyacrylamide gel. The same primer used for primer extension was also used to generate a sequencing ladder using a PCR-generated template. The samples were electrophoresed at 35 W for 1.75 hours; fixed in 10% methanol, 10% acetic acid; dried; and exposed to film overnight at room temperature.

II.9 WESTERN ANALYSIS

Cell free extracts of *S. clavuligerus* were prepared according to Ingram *et al.*. (1989). Mycelia grown in liquid culture for 24-48 hours were collected by filtration through Whatman No.1 filter paper and washed with 10 mL of TDE buffer (0.05 M Tris; 0.01 mM EDTA; 0.1 mM DTT; pH 7.2). Mycelia were scraped into a small beaker and resuspended in 10 mL of TDE buffer and dispensed into 1.5 mL Eppendorf tubes. Cell free extracts were prepared by sonication for 3 × 15 seconds at setting 3 of a Branson Sonifier 450 with a 2.5 mm diameter probe. Cell debris was removed by centrifugation at 15,000 rpm in an Eppendorf 5414 C centrifuge for 15 minutes at 4°C. Total protein was measured by the method of Bradford (1976) using the Bio-Rad microassay procedure described by the supplier. Bovine gamma gobulin was used as the protein standard. Supernatants were frozen at -70°C. Samples (10 μg) were suspended in 5 μL of loading dye (0.125 M Tris, pH 6.8; 6% SDS; 30% glycerol; 15% 2-mercaptoethanol; 0.003% bromophenol blue) and PBS buffer (1.5 mM NaH_2PO_4, 8.1 mM Na_2HPO_4, 0.145 M NaCl) to a total volume of 15 μL. The samples were denatured by heating at 80°C for 10 minutes. The samples were separated on a 10% SDS-PAGE gel with a 3.2% polyacrylamide (37.5:1 acrylamide: bisacrylamide) 0.068% SDS stacking gel with running buffer (0.05 M Tris, 0.38 M Glycine, 0.1% SDS). Gels were run at 40 mA for

5 hours. Transfer to a PVDF membrane was carried out at 4°C and 52V overnight using a Bio-Rad Transblot apparatus (Richmond, California). The membrane was washed 3 × 5 minutes in milli-Q H_2O, then for 5 minutes in TBS wash buffer (0.02 M Tris; 0.14 M NaCl; 0.1% Tween 20; pH 7.6). The membrane was blocked in 40 mL of blocking buffer (0.02 M Tris; 0.14 M NaCl; 0.1% Tween 20; 5% BSA; pH 7.6) for 1 hour at 25°C. The membrane was reacted with a 1:5000 dilution of primary antibody (rabbit α-CcaR) in 20 mL of TBS for 1 hour at 25°C, then washed in TBS wash buffer for 2 × 5 minutes and 1 × 15 minutes. The membrane was then reacted with a 1:5000 dilution of secondary antibody (donkey α-rabbit Ig horseradish peroxidase) in 20 mL of TBS for 1 hour at 25°C, then washed in TBS wash buffer for 2 × 5 minutes and 1 × 15 minutes. The reactions were carried out with gentle rocking on a rocker platform from Bellco Biotechnology, and the washes were done on an orbital shaker from New Brunswick Scientific Co. Proteins were detected using ECL (enhanced chemiluminescence) Western system reagents and protocol.

II.10 REPORTER GENE ANALYSIS

II.10.1 Promoter probe analysis using the *xylE* reporter gene

The high copy number *Streptomyces* plasmid pIJ4083 (Clayton and Bibb, 1990) was used to identify *S. clavuligerus bldA* promoters. The plasmid contains the promoter-less *xylE* reporter gene (Zukowski *et al.*, 1983) which encodes the enzyme catechol-2,3-dioxygenase (C230). A multiple cloning site is located immediately upstream of the *xylE* gene and was used to directionally clone the putative promoter fragments. The constitutively expressed *ermE* * promoter (Bibb *et al.*, 1985; Bibb *et al.*, 1986) was used as a positive control and the promoter-less vector was used as a negative control.

S. lividans cultures containing the recombinant plasmids were maintained on R2YE agar plates containing 50 µg/mL thiostrepton and as frozen stocks. *S. clavuligerus* cultures containing the recombinant plasmids were maintained on MYM + 5 µg/mL of thiostrepton and as frozen stocks. Visual assessment of catechol dioxygenase activity of *S. clavuligerus* cultures containing the recombinant plasmids was performed by spraying the plates containing transformants with a 0.5 M solution of catechol and looking for colonies that turned yellow. Visual assessment of catechol dioxygenase activity was performed on minimal medium (0.05% L-asparagine, 2.8 mM K_2HPO_4, 0.8 mM $MgSO_4$, 0.036 mM $FeSO_4$, 2% agar) + 0.5 % maltose, + 5 µg/mL thiostrepton, as some of the transformants produced a yellow-colored pigment (possibly holomycin) on rich media. The yellow color was suppressed on minimal medium.

II.10.2 Catechol dioxygenase assays of *S. lividans* TK24 and *S. clavuligerus* NRRL 3585 transformants containing recombinant pIJ4083 plasmids

Cell free extracts of *S. lividans* and *S. clavuligerus* containing the desired plasmids were prepared according to Ingram *et al.* (1989). Mycelia grown in liquid culture for 36 hours were collected by filtration through Whatman No.1 filter paper and washed with 50 mL of wash buffer (20 mM potassium salt-phosphate buffer, pH 7.2). Mycelia were scraped into a small beaker and resuspended in 3 mL of sample buffer (100 mM KPO_4, pH 7.5; 20 mM EDTA; and 10% (v/v) acetone) and dispensed into 1.5 mL Eppendorf tubes. Cell free extracts were prepared by sonication for 3 × 10 seconds at setting 2-5 of a Branson Sonifier 450 with a 2.5 mm diameter probe. Cell debris was removed by centrifugation at 15,000 rpm in an Eppendorf 5415 C centrifuge for 15 minutes at 4°C. Supernatants were either put on ice to be assayed immediately or frozen at $^-70$°C.

Total protein was measured by the method of Bradford (1976) using the Bio-Rad microassay procedure described by the supplier. Bovine gamma gobulin was used as the protein standard. Catechol dioxygenase activity was measured as described by Zukowski *et al.* (1983). Up to 100 µL of cell free extract was added to 2.9 mL of prewarmed reaction mixture (100 mM potassium salt-phosphate buffer, pH 6.8 and 10 µL of 20 mM catechol in ethanol) to a final volume of 3 mL. The reaction was mixed and the change in absorbance at 375 nm was measured at 1 minute intervals for 10 minutes at 30°C. The amount of cell-free extract used was adjusted to give a linear change in absorbance with the maximum absorbance falling below 1.0. In some cases it was necessary to dilute the cell free extract, in which case the extract was diluted in reaction buffer. Under these conditions the molar absorption coefficient of 2-hydroxymuconic semialdehyde is reported as 3.3×10^4 (Sala-Trepat and Evans, 1971). One milliunit (mU) is defined as the formation of 1 nmol of 2-hydroxymuconic semialdehyde per minute at 30°C. The activity is reported as mU/mg of total protein.

II.10.3 Green Fluorescent Protein reporter gene analysis

The *gfp-mut3* gene (Cormack *et al.*, 1996), containing three TTA codons was cloned into the high copy number *Streptomyces* vector pIJ486, under the control of the constitutively expressed *ermE** promoter (Markus and Leskiw, 1997). This vector was introduced into *S. lividans* and *S. clavuligerus* NRRL 3585 (wild-tpe) and *bldA* strains and the colonies were assessed for GFP under a microscope using a mercury lamp and fluorescein isothiocyanate filter.

II.11 ANTIBIOTIC BIOASSAYS

Antibiotic bioassays were performed by the agar diffusion method to determine the presence of cephamycin C or clavulanic acid. Antibiotic production from *S. clavuligerus* cultures grown on the surface of cellophane disks was assayed by removing a plug of agar from under the cellophane disc with a No.4 cork borer (7 mm diameter) and placing it on the surface of MYM (1.5% agar) plates. The antibiotic was allowed to diffuse from the plug into the MYM agar for 1 hour, the plug was removed and then the MYM plate was overlaid with 5 mL of soft MYM agar (0.6%) supplemented with an aliquot of the appropriate indicator bacterium (see below). For antibiotic assays of *S. clavuligerus* grown in liquid cultures, the MYM agar plate was first overlaid with soft agar and then sterile 1 cm filter paper disks were placed on the agar and 100 µL of culture supernatant was added to the disk. To determine the presence of cephamycin C, the soft agar was inoculated with 100 µL of a glycerol stock of *E. coli* ESS. To determine the presence of clavulanic acid the soft agar was inoculated with 20 µL of *S. aureus* N2 glycerol stock and 3 µg/mL of penicillin G. The penicillin G was omitted from the control plates for the *S. clavuligerus* antibiotic assay. All plates were incubated overnight at 37°C and the size of the zone of inhibition was measured. For the clavulanic acid bioassay, the control plates indicated the presence of the antibiotics holomycin, tunicamycin, and cephamycin C and an increase in the diameter of the zone of inhibition on the test plates indicated the presence of clavulanic acid. The antibiotic bioassay for the *S. clavuligerus* liquid time course was performed on large square agar plates. The plates were UV-sterilized and filled with 325 mL of 1.5 % MYM agar. The cephamycin C test plates were overlaid with 1 mL of *E. coli* ESS in 50 mL of soft agar. The clavulanic acid plates were overlaid with 200 µL of *S. aureus* N2 +/- 3µg/mL of penicillin G in 50 mL of soft agar.

III. RESULTS

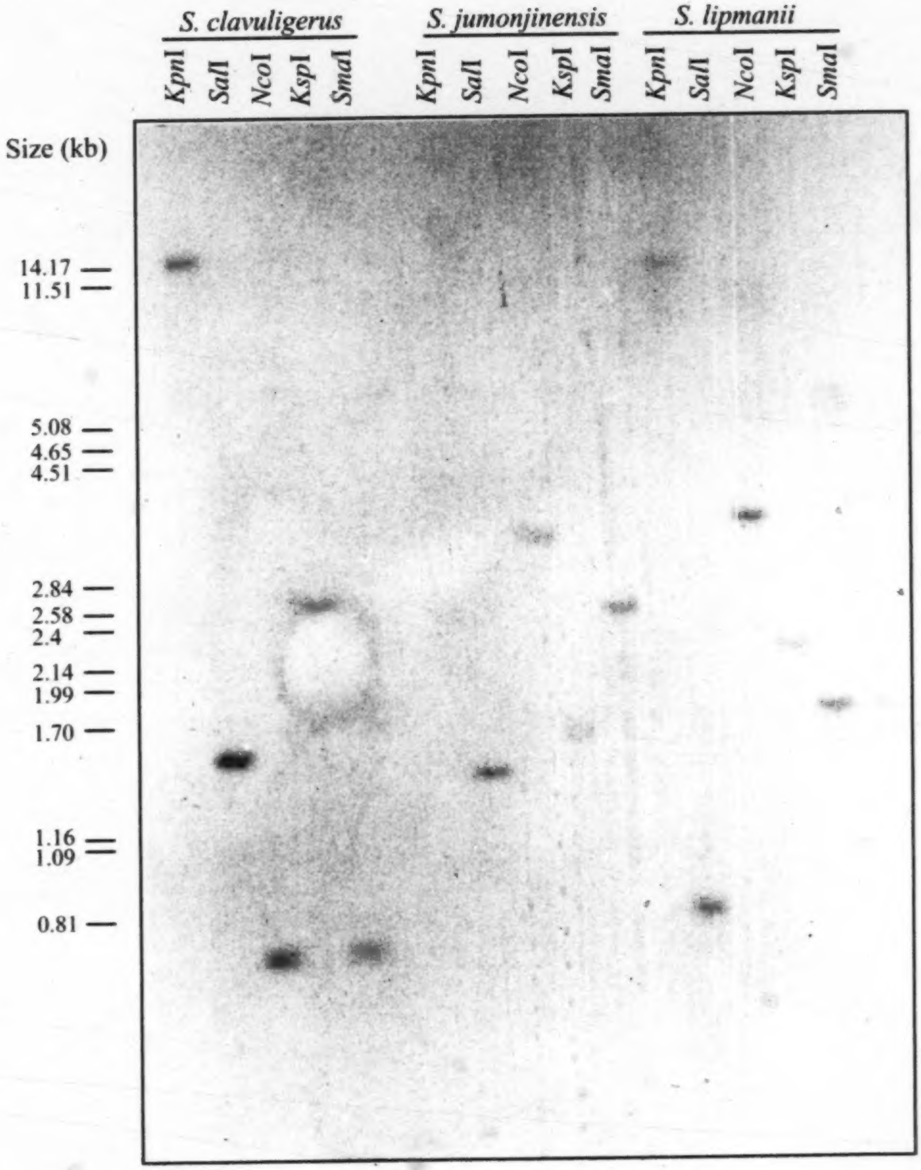
Although the *bldA* gene has been identified in *S. coelicolor, S. lividans*, and *S. griseus*, it has only been extensively characterized in *S. coelicolor*. We decided to identify and characterize the *bldA* homologue in an additional *Streptomyces* strain to determine if the phenotype of the mutant would prove to be sporulation and antibiotic deficient and provide some evidence for the role of the *bldA* gene in *Streptomyces* as a whole. We decided to study *Streptomyces* strains that produced a variety of β-lactam antibiotics, as the *bldA* gene had not previously been identified in this group (although some *S. griseus* strains produces the β-lactams cephamycin A and B, the strain that the *bldA* gene has been identified in has not been characterized as a β-lactam producer). Also, the β-lactam producers are not closely related to *S. coelicolor* and any similarities that are found with the *bldA* phenotype would likely reflect a common role of the *bldA* gene in *Streptomyces.*

III.1 INITIAL SCREENING OF β-LACTAM PRODUCING *STREPTOMYCES* STRAINS

To identify a *bldA* homologue in a β-lactam producing streptomycete, three β–lactam producing strains were initially screened. Chromosomal DNA from *S. clavuligerus, S. jumonjinensis,* and *S. lipmanii* was digested with *Kpn*I, *Sal*I, *Nco*I, *Ksp*I *(Sac*II), and *Sma*I. The DNA was separated on a 1% agarose gel and transferred to a nylon membrane. The membrane was probed with a ^{32}P-dATP random primer-labelled, 286 bp PCR-amplified fragment of the *S. coelicolor bldA* gene. Hybridizations and washes were carried out at low stringency (37°C) to allow hybridization to sequences with up to 25% mismatched bases. Under these conditions only one band was observed for each digest (Figure III.1.1), which most likely indicates that a single-copy *bldA*

Figure III.1.1

Hybridization of *S. coelicolor bldA* to *S. clavuligerus, S. jumonjinensis,* and *S. lipmanii* DNA. Two micrograms of chromosomal DNA of *S. clavuligerus, S. jumonjinensis,* and *S. lipmanii* were digested with *Kpn*I, *Sal*I, *Nco*I, *Ksp*I *(Sac*II), and *Sma*I. The DNA was separated on a 1% agarose gel and transferred to a nylon membrane. The membrane was probed with a ^{32}P-labelled DNA fragment containing the *S. coelicolor bldA* gene. The 286 bp probe was generated by PCR of *S. coelicolor* chromosomal DNA using the primers BKL2 (5'-CATGGATCCACCCGGTAACTGATG-3'), homologous to a region upstream from the *bldA* gene, and BKL15 (5'-GCCGCTGAGTCGCAACCA-3'), homologous to a region downstream of the *bldA* gene. Hybridizations and washes were carried out at low stringency (37°C). The DNA size standard was λ DNA digested with *Pst*I.


S. clavuligerus
S. jumonjinensis
S. lipmanii
KpnI
SalI
NcoI
KspI
SmaI
KpnI
SalI
NcoI
KspI
SmaI
KpnI
SalI
NcoI
KspI
SmaI
Size (kb)
14.17
11.51
5.08
4.65
4.51
2.84
2.58
2.4
2.14
1.99
1.70
1.16
1.09
0.81

homologue is present in each of the three *Streptomyces* species tested. While it is possible that the probe is hybridizing to a leucyl-tRNA other than *bldA*, this possibility is less likely since a comparison of *E. coli* leucyl-tRNAs show that there is significant divergence between $tRNA^{Leu}_{UAA}$ and other leucyl-specifying tRNAs, so it would follow that the sequences most homologous to the *S. coelicolor bldA* gene will be $tRNA^{Leu}_{UAA}$ genes.

III.2 CLONING OF THE *S. clavuligerus bldA* GENE

While *bldA* homologues were identified in all three *Streptomyces* species tested, a decision was made to continue study of only one species. The species *S. clavuligerus* was chosen for further study since it is currently the focus of research by a number of laboratories and much information has been elucidated regarding production of the β-lactam antibiotic cephamycin C and the β-lactamase inhibitor clavulanic acid.

Initial efforts to clone the *bldA* homologue from chromosomal restriction digests were unsuccessful so subsequent efforts were made to subclone the *bldA* homologue from an *S. clavuligerus* cosmid library. The cosmid library had been generated by ligating chromosomal DNA, partially digested with *Sau3*AI, into the *Bam*HI digested pLAFR3 vector and cloning into *E. coli* VCS257 (Doran *et al.*, 1990). Membranes for colony hybridization were obtained from Dylan Alexander, University of Alberta, Edmonton, Alberta. These membranes were probed with the ^{32}P-dATP random primer-labelled 286 bp PCR-amplified fragment of the *S. coelicolor bldA* gene under the same conditions used to detect the *bldA* homologues in chromosomal digests. A single hybridizing clone was detected which corresponded to a clone in microplate 11. Twenty one clones from plate 11 were cultured and used to generate a dot-blot to further identify the single hybridizing clone. Clone pLAFR3-11C1, was identified as the only clone from the cosmid library to hybridize to the *S. coelicolor bldA* gene. In order to verify that the probe hybridized to the insert in this cosmid, and to identify a suitable fragment for subcloning, cosmid DNA was isolated, digested with a number of restriction enzymes, separated on an agarose gel, transferred to a nylon membrane and probed with the *S. coelicolor bldA* gene. Three hybridizing fragments, 1.5 kb *Sal*I, 0.7 kb *Sma*I, and 2.6 kb *Ksp*I, were detected which corresponded to similar sized fragments identified in chromosomal digests of *S. clavuligerus* (Figure III.2.1). To subclone the 1.5 kb *Sal*I fragment, pLAFR3-11C1 was digested with *Sal*I, the fragments were separated on an ultrapure 1% agarose gel, stained in ethidium bromide and the two bands in the 1.5 kb range were excised. The DNA was purified using Gene Clean™, ligated into *Sal*I-digested pUC119 and transformed into *E. coli* DH5α competent cells. The transformants were plated on LB + ampicillin agar containing Xgal and IPTG. The transformants containing inserts were screened by colony hybridization with the *S. coelicolor bldA* probe. One of the transformants that hybridized to the *bldA* probe and contained a 1.5 kb insert was chosen for further study. This clone was named p9S+.

To ensure that the *S. clavuligerus* DNA had not undergone any rearrangements during cloning into the pLAFR cosmid, the 1.5 kb *Sal*I insert was purified, random primer labelled with ^{32}P-dATP and used to probe *S. clavuligerus* chromosomal digests (Figure III.2.2). Hybridization and washes were carried out at 65°C. The probe showed the same pattern of hybridization that was previously observed with the *S. coelicolor bldA* probe. Some additional hybridizing bands are present in the *Nco*I, *Sma*I, *and Ksp*I digestions which are presumably due to the presence of these restriction sites within the 1.5 kb *Sal*I region.

Figure III.2.1

Hybridization of the *S. coelicolor bldA* gene to pLAFR-11C1 cosmid DNA digested with a variety of restriction endonucleases. pLAFR-11C1 DNA was digested with *Eco*RI+*Hin*dIII (to excise the *S. clavuligerus* insert), *Bgl*II, *Kpn*I, *Ksp*I (*Sac*II), *Sal*I, and *Sma*I. The digests were separated on a 1% agarose gel, transferred to a nylon membrane and probed with a 286 bp ^{32}P-dATP random primer-labelled PCR-amplified fragment of the *S. coelicolor bldA* gene. The 286 bp probe was generated as described previously (Figure III.1.1). Hybridizations and washes were carried out at 37°C. The additional hybridizing bands in the *Ksp*I digestions are presumably due to incomplete digestion of the cosmid DNA. The molecular weight markers (*Pst*I-digested λ DNA) were detected by probing with random-primer, ^{32}P-labelled λ DNA. In this case, hybridization and washes were carried out at 65°C.

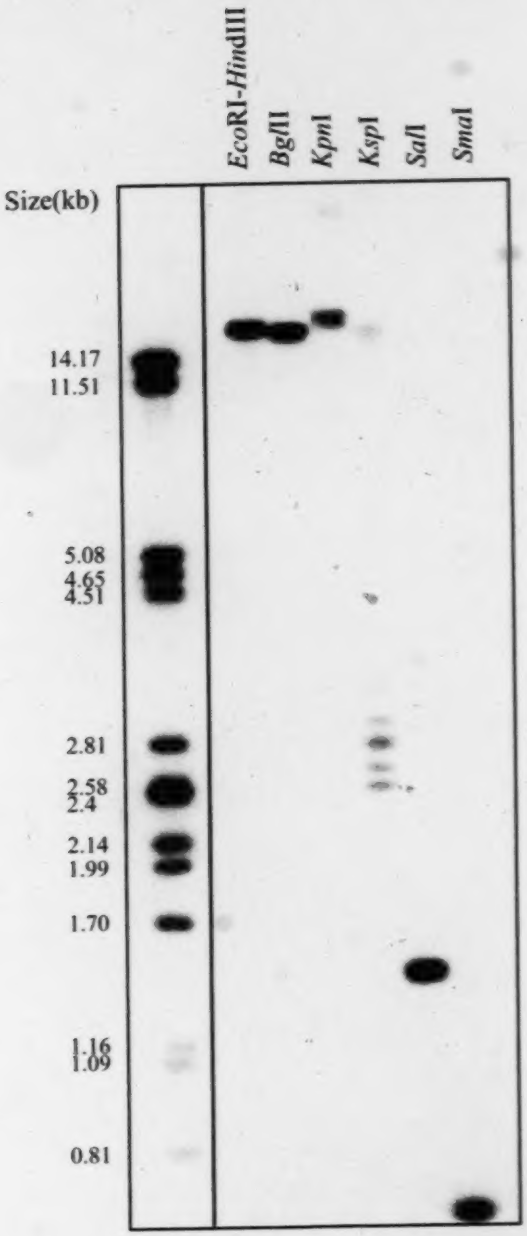
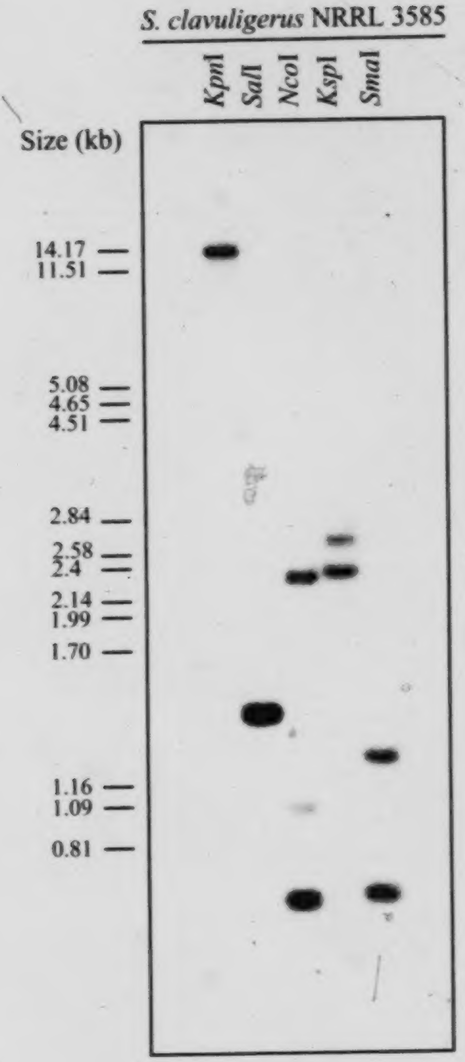

Size(kb)
14.17
11.51
5.08
4.65
4.51
2.81
2.58
2.4
2.14
1.99
1.70
1.16
1.09
0.81
EcoRI-HindIII
BglII
KpnI
KspI
SalI
SmaI

Figure III.2.2

Hybridization of the *S. clavuligerus* 1.5 kb *Sal*I restriction fragment to *S. clavuligerus* chromosomal DNA. *S. clavuligerus* DNA was digested with *Kpn*I, *Sal*I, *Nco*I, *Ksp*I, and *Sma*I, separated on a 1% agarose gel, transferred to a nylon membrane and probed with the ^{32}P-dATP random primer-labelled 1.5 kb *Sal*I fragment which was cloned from cosmid pLAFR11C1. Hybridization and washes were carried out at 65 °C. The DNA size standard was λ DNA digested with *Pst*I.


S. clavuligerus NRRL 3585
KpnI
SalI
NcoI
KspI
SmaI
Size (kb)
14.17
11.51
5.08
4.65
4.51
2.84
2.58
2.4
2.14
1.99
1.70
1.16
1.09
0.81

III.3 SUBCLONING OF THE *S. clavuligerus bldA* GENE

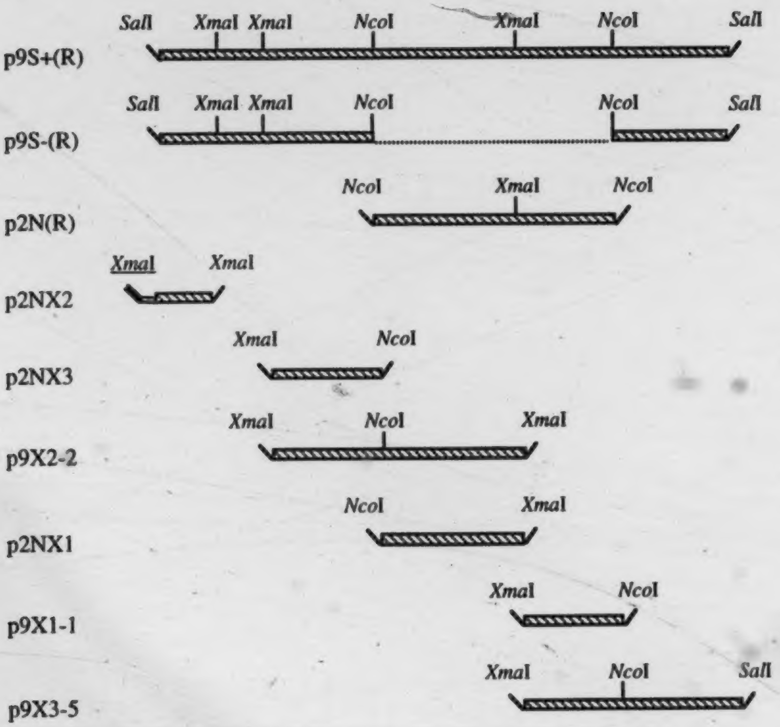
For the purpose of generating sequencing templates, the insert in p9S+ was subcloned using *Nco*I, *Xma*I (*Xma*I recognizes the same nucleotide sequence as *Sma*I but generates 5' overhangs instead of blunt ends), and *Sal*I restriction enzymes alone or in combination. The digests were separated on 5% polyacrylamide gels and the liberated inserts were purified by the crush and soak procedure (described in Section II.6.2). The purified inserts were cloned into pUC119 (p9 clones) or pUC120 (p2 clones). pUC120 was used to subclone inserts with an *Nco*I end as this plasmid contains an *Nco*I site in the multiple cloning site. The host used was *E. coli* MV1193 which contains an F pilus and can be used to generate single-stranded templates for sequencing using a helper phage. The original clone, p9S+, was also transformed into MV1193 prior to sequencing. The subclones generated are listed in Figure III.3.1.

III.4 SEQUENCE ANALYSIS OF THE 1.5 kb *Sal*I FRAGMENT

Single-stranded sequencing templates of the clones listed in Figure III.3.1 were generated using the helper phage M13K07, and sequenced using the universal primer. Synthetic oligonucleotide primers were generated to sequence the gaps (listed in Table II.1.1) and both strands were sequenced for the entire region. One region that contained a series of repetitive units and was over 80% G+C could not be sequenced with Sequenase™ and was sequenced by a sequencing service. The sequence of the 1.5 kb fragment is shown in Figure III.4.1. The sequence of the *bldA* $\text{tRNA}^{\text{Leu}}_{\text{UAA}}$ gene that corresponds to the mature processed form of the tRNA (tDNA) is located from bases 681-764. The *bldA* gene is presumably transcribed as a monocistronic transcript as no other tRNA genes are located in this region. Comparison of the putative *S. clavuligerus*

Figure III.3.1

Restriction map of the 1.5 kb *S. clavuligerus Sal*I insert in p9S+. Subclones were generated by digesting the 1.5 kb insert with a combination of *Nco*I, *Xma*I, and *Sal*I . The fragments were gel purified and subcloned into either pUC119 or pUC120. The multiple cloning site of pUC119 and pUC120 is shown and the direction of sequencing with the universal primer is indicated by an arrow. The brackets around the *Nco*I site indicate that this site is present in the multiple cloning site of pUC120 but not pUC119. Subclones generated in pUC119 are indicated as p9 clones and those generated in pUC120 are indicated as p2 clones. Some inserts were cloned in both orientations, these are indicated by (R). The duplicated vector sequence of p2NX2 is indicated by a double line and underlined restriction site. The *S. clavuligerus* sequence is indicated by hatched boxes.


SalI
XmaI
XmaI
NcoI
XmaI
NcoI
SalI
p9S+(R)
SalI
XmaI
XmaI
NcoI
NcoI
SalI
p9S-(R)
NcoI
XmaI
NcoI
p2N(R)
XmaI
XmaI
p2NX2
XmaI
NcoI
p2NX3
XmaI
NcoI
XmaI
p9X2-2
NcoI
XmaI
p2NX1
XmaI
NcoI
p9X1-1
XmaI
NcoI
SalI
p9X3-5

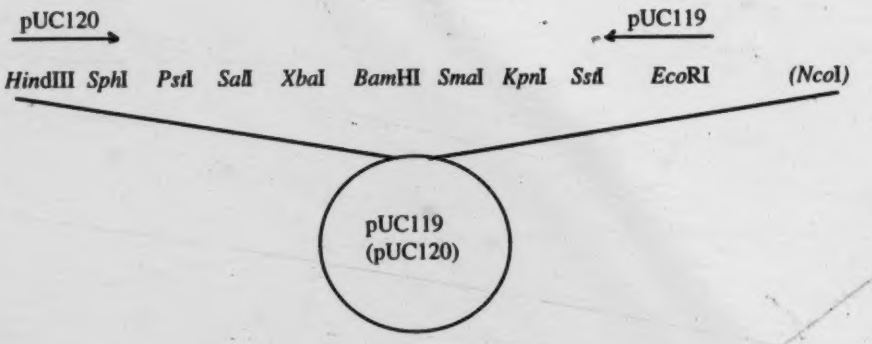

pUC120
pUC119
HindIII
SphI
PstI
SalI
XbaI
BamHI
SmaI
KpnI
SstI
EcoRI
(NcoI)
pUC119
(pUC120)

Figure III.4.1

The sequence of the *S. clavuligerus* 1.5 kb *Sal*I fragment containing the *bldA* tRNA gene. The *Sal*I, *Sma*I, *Nco*I, *and Ksp*I recognition sites are indicated. The boxed sequence corresponds to the *bldA* $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{UAA}}$ gene (tDNA sequence) and is located from bases 681-764 (the anticodon is marked by *). The two putative open reading frames are indicated as well as the corresponding amino acid sequences.

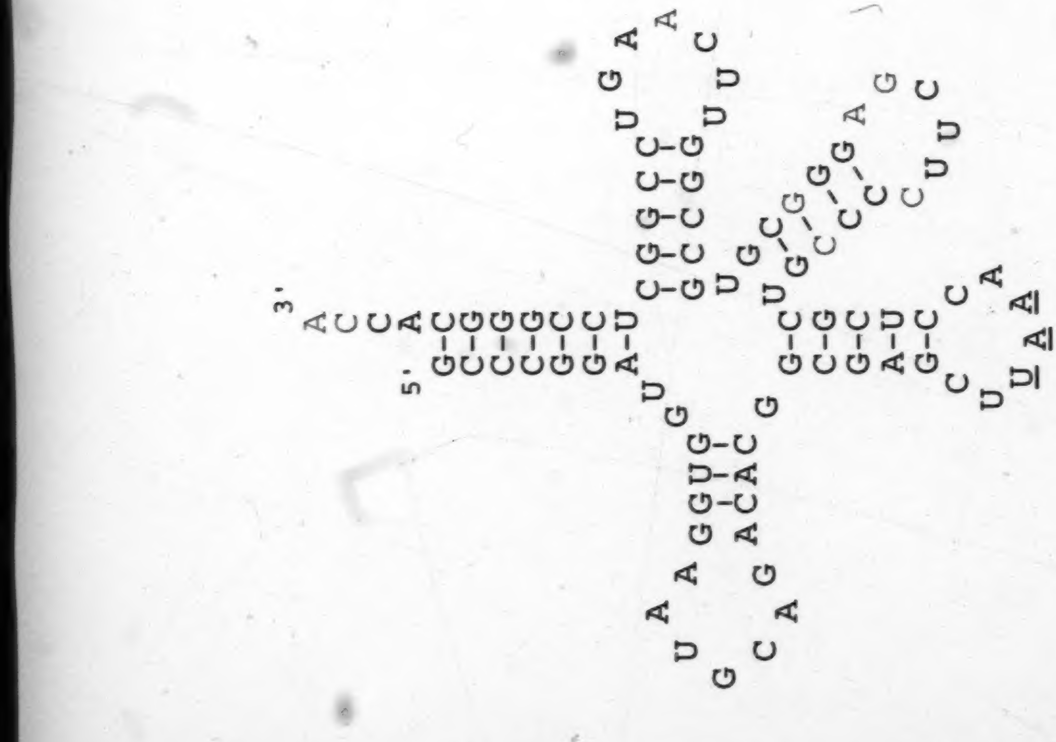
```
 SalI
GTCGACGCTGACCCGGTCCAGGGCCACCACCTGGGTCTCGCCCTCCCCGTACACCTTGGA
CAGCTGCGACTGGGCCAGGTCCCGGTGGTGGACCCAGAGCGGGAGGGGCATGTGGAACCT   60
 D  V  S  V  R  D  L  A  V  V  Q  T  E  G  E  G  Y  V  K  S

CAGTTCCGTGGCGCGGGCGGCCACGGCGGTGGCGCGGTGTGCGATGGGCATGGTGGTCAC
GTCAAGGCACCGCGCCCGCCGGTGCCGCCACCGCGCCACACGCTACCCGTACCACCAGTG   120
 L  E  T  A  R  A  A  V  A  Y  A  R  H  A  I  P  M  T  T fM ORF1
                                     SmaI
GGGGGCACTCCTGTTCGGGCGGGTCGATCCGGGGACCCGGGGGCGGCCACTGCGGGAGGC
CCCCCGTGAGGACAAGCCCGCCCAGCTAGGCCCCTGGGCCCCCGCCGGTGACGCCCTCCG   180

CGTCTCCGAGGACACACACCATCCTGTCCTTGTTCCGGGCGGGGATCGTCAGCCCCCGTG
GCAGAGGCTCCTGTGTGTGGTAGGACAGGAACAAGGCCCGCCCCTAGCAGTCGGGGGCAC   240

        SmaI
CCCGTTCCCGGGGCCCCCTTTGGTCGCACCAGGAGCGGCGATCATCCTCCTTGGGTATGA
GGGCAAGGGCCCCGGGGGAAACCAGCGTGGTCCTCGCCGCTAGTAGGAGGAACCCATACT   300

CGGTGCCCCTGAGGGCGGAAAGGGGACGGGGCGCTCCGGAGCGTCGTCACTCTTCGCTGC
GCCACGGGGACTCCCGCCTTTCCCCTGCCCCGCGAGGCCTCGCAGCAGTGAGAAGCGACG   360

GGTGTCCGGGGGCGCGGGCGCCGCCCTCGATCCGGTGAGCGTCGCGTGGCTGTTCGGTGG
CCACAGGCCCCCGCGCCCGCGGCGGGAGCTAGGCCACTCGCAGCGCACCGACAAGCCACC   420

AGGCCGGGTGCCGACCGCGTGTGCGGGGTGCGGAGAGAGGGCGGACGCGTCCGAAGCCGA
TCCGGCCCACGGCTGGCGCACACGCCCCACGCCTCTCTCCCGCCTGCGCAGGCTTCGGCT   480
                                                        NcoI
CTCGGTTGCGACGAGTTTCCGTCATTCGCGTGCGGGGGTCGTGACGCCCGGCGTTCGCCA
GAGCCAACGCTGCTCAAAGGCAGTAAGCGCACGCCCCCAGCACTGCGGGCCGCAAGCGGT   540

TGGAACGCCTTGTGATCCTCTTCCACCTGGGCTGACGCACCCTCAAGCGGCAATAAAATA
ACCTTGCGGAACACTAGGAGAAGGTGGACCCGACTGCGTGGGAGTTCGCCGTTATTTTAT   600

AGACAACATCGGGCCACGGGCCGCCGGGCGAGTCGATCTCCCCGGATAGGCTCATGTGCT
TCTGTTGTAGCCCGGTGCCCGGCGGCCCGCTCAGCTAGAGGGGCCTATCCGAGTACACGA   660
                     bldA tDNA                        ***
TACGCGGAGCCACGCTCCTGGCCCGGATGGTGGAATGCAGACACGGCGAGCTTAAACCTC
ATGCGCCTCGGTGCGAGGACCGGGCCTACCACCTTACGTCTGTGCCGCTCGAATTTGGAG   720

GCTGGCCTTCATGGCCGTGCCGGTTCGAGTCCGGCTCCGGGCACTCTCCGTAACGAGAGC
CGACCGGAAGTACCGGCACGGCCAAGCTCAGGCCGAGGCCCGTGAGAGGCATTGCTCTCG   780
```

```
                                                     KspI
                                                     ______
GGGACGCCGGGTAGTTGAATGCTTGACATTCAAGATCGCCTCATCGTCGCACCCCGCGGC  840
CCCTGCGGCCCATCAACTTACGAACTGTAAGTTCTAGCGGAGTAGCAGCGTGGGGCGCCG

CCCCGCGCGCCACCCCTGACGAGCCCCCGATCCCCCCTGATCCACCCCTGATCCACCCCC  900
GGGGCGCGCGGTGGGGACTGCTCGGGGGCTAGGGGGGACTAGGTGGGGACTAGGTGGGGG
                      SmaI
                      ______
GATCCGGCCCCGAGACCCGCGCCCCGGGATCTCCGCCCCCCTCCTCGTGACGCATCCGAG  960
CTAGGCCGGGGCTCTGGGCGCGGGGCCCTAGAGGCGGGGGGAGGAGCACTGCGTAGGCTC
                  KspI
                  ______
GAGACACCCCGGCCGCTCCCCGCGGCCCCGGACACCCCGGCCGCCCGTCCGGCCGAACGG  1020
CTCTGTGGGGCCGGCGAGGGGCGCCGGGGCCTGTGGGGCCGGCGGGCAGGCCGGCTTGCC

CGACACGCCCCACCCGGACCCGTCCGCGACCGGCGCCCCGCGTCCGGGCAGGCGGCATGG  1080
GCTGTGCGGGGTGGGCCTGGGCAGGCGCTGGCCGCGGGGCGCAGGCCCGTCCGCCGTACC

CCGGAAACGTATGCCAGGTCACAATCGGGAAAGATCCTCCCTCAGCACTAGGCTTCTTCC  1140
GGCCTTTGCATACGGTCCAGTGTTAGCCCTTTCTAGGAGGGAGTCGTGATCCGAAGAAGG
                                          NcoI     ORF2  fM
                                          ______
TTTGCGCCCCTATTACTCTTGACGCAAGGCCGCGCACGGCGGCCATGGAGGAGTGAGATG  1200
AAACGCGGGGATAATGAGAACTGCGTTCCGGCGCGTGCCGCCGGTACCTCCTCACTCTAC

R  S  S  N  P  V  F  S  R  R  G  F  S  R  D  N  G  I  A  G
AGGAGCAGTAACCCGGTCTTCTCGCGACGGGTTTCAGCCGCGACAACGGCATTGCGGGC  1260
TCCTCGTCATTGGGCCAGAAGAGCGCTGCCCCAAAGTCGGCGCTGTTGCCGTAACGCCCG

F  N  A  Q  Q  Q  A  G  S  P  V  A  G  N  P  Y  A  Q  G  A
TTCAACGCGCAGCAGCAGGCCGGGAGCCCTGTCGCCGGTAACCCCTATGCCCAGGGCGCG  1320
AAGTTGCGCGTCGTCGTCCGGCCCTCGGGACAGCGGCCATTGGGGATACGGGTCCCGCGC

A  N  P  Y  A  T  N  P  Y  A  P  A  D  T  Q  L  G  A  A  Q
GCCAACCCGTACGCCACCAACCCGTACGCACCTGCCGACACCCAGCTCGGCGCGGCCCAG  1380
CGGTTGGGCATGCGGTGGTTGGGCATGCGTGGACGGCTGTGGGTCGAGCCGCGCCGGGTC

A  P  A  H  R  V  M  T  I  D  D  V  V  A  R  T  A  M  T  L
GCCCCCGCCCACCGGGTGATGACGATCGACGACGTCGTGGCGCGTACCGCGATGACGCTC  1440
CGGGGGCGGGTGGCCCACTACTGCTAGCTGCTGCAGCACCGCGCATGGCGCTACTGCGAG

G  T  V  V  V  T  A  M  L  A  W  F  L  L  P  V  D
GGCACGGTCGTCGTCACCGCGATGCTCGCGTGGTTCCTGCTGCCCGTCGAC  1491
CCGTGCCAGCAGCAGTGGCGCTACGAGCGCACCAAGGACGACGGGCAGCTG
                                             ______
                                             SalI
```

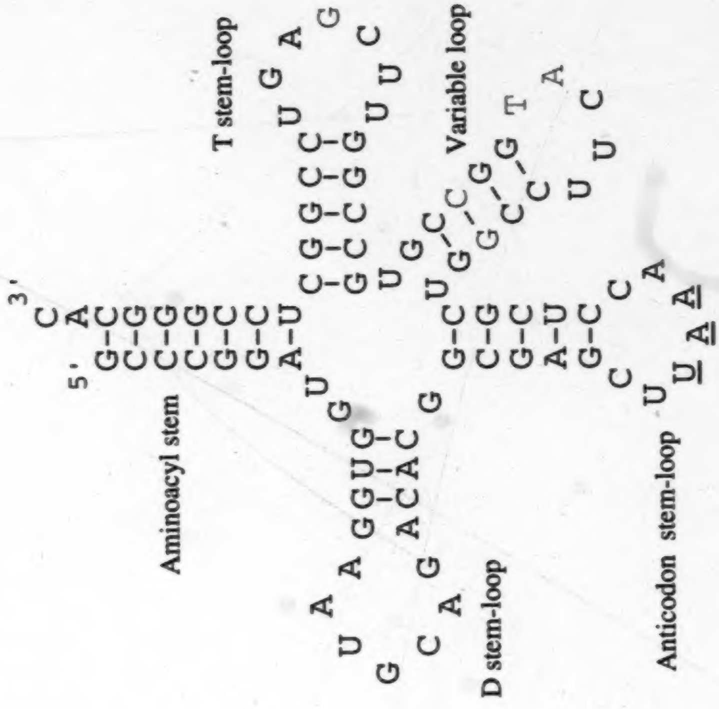
bldA tRNA with the *S. coelicolor bldA* tRNA showed 93% similarity between the two tRNAs with a single base gap (Figure III.4.2). Most of the sequence differences are confined to the variable loop, including the single base gap. The differences between the two variable stems are compensatory, the *S. coelicolor* stem has a CG pairing while the *S. clavuligerus* stem has a GC pairing. There is a single base difference in the T–loop at position 57 although both tRNAs contain a purine which is typically conserved. The *S. clavuligerus bldA* tDNA is 83-4 nucleotides compared to the *S. coelicolor bldA* tDNA which is 87 nucleotides. The differences in length are due to an extra base in the *S. coelicolor* variable loop and the gene-encoded 3' CCA terminus. All mature tRNAs contain the nucleotides CCA at the 3' end, and in most prokaryotes this is gene-encoded, however, the majority of *Streptomyces* tRNA genes encode only the first cytidine (C74) of the C74-C75-A76 sequence. Thus the C75-A76 nucleotides (or CCA nucleotides) must be added post-transcriptionally to the majority of *Streptomyces* tRNAs. There are some tRNA genes in *Bacillus subtilis* and in the *E. coli* phage T4 that lack the CCA end, and both *E. coli* and *B. subtilis* contain enzymes that can add the CCA end (Vold, 1985), so the post-transcriptional addition of these nucleotides is not without precedent among prokaryotes. There are numerous bases other than the 3' CCA that are conserved in most tRNA genes, and depending on their degree of conservation are called invariant or semi-invariant bases. However, as more tRNA genes are sequenced it is apparent that none of the bases are truly invariant and an invariant or semi-invariant base in one species may not be conserved in another (Dirheimer *et al.*, 1995). In order to first identify which bases are invariant and semi-invariant in *Streptomyces* tRNAs, all of the available sequence of *Streptomyces* tRNA genes were compared (Table III.4.1). From the

Figure III.4.2

The deduced secondary structure of the *S. clavuligerus*, and *S. coelicolor bldA* tRNAs. The base differences between the two tRNAs are indicated in grey, the anticodon is underlined. The aminoacyl stem, anticodon stem-loop, variable loop, and Ŧ (formerly TΨC) stem-loop of the *S. clavuligerus bldA* tRNA are labelled.



S. coelicolor bldA


5'
3'
T stem-loop
Aminoacyl stem
Variable loop
D stem-loop
Anticodon stem-loop

S. clavuligerus bldA

Table III.4.1 Comparison of *Streptomyces* tDNA Gene Sequences to Identify Invariant and Semi-Invariant Bases

Organism	tRNA	Anti-codon	Aminoacyl stem (5')		D stem-loop[2]				Anticodon stem-loop		
			1-7	8-9	10-13	14-21	22-25	26	27-31	32-38	39-43
S. clavuligerus	Leu	TAA	GCCCGGA	TG	GTGG	AAT GCAG	ACAC	G	GCGAG	CT TAA AC	CTCGC
S. coelicolor	Leu	TAA	GCCCGGA	TG	GTGG	AAT GCAG	ACAC	G	GCGAG	CT TAA AC	CTCGC
S. lividans	Leu	TAA	GCCCGGA	TG	GTGG	AAT GCAG	ACAC	G	GCGAG	CT TAA AC	CTCGC
S. griseus	Leu	TAA	GCCCGGA	TG	GTGG	AAC GTAG	ACAC	G	GCGAG	CT TAA AC	CTCGC
S. coelicolor	leuUα	GAG	GTCCGGG	TG	GCGG	AAT GGCAG	ACGC	G	CTAGC	TT GAG GT	GCTAG
S. coelicolor	leuUβ	GAG	GTCCGGG	TG	GCGG	AAT GGCAG	ACGC	G	CTAGC	TT GAG GT	GCTAG
S. lividans	Ser	GCT	GGAGGCG	TC	GCCT	AGTCCGGTCTA	TGGC	G	CCGCA	CT GCT AA	TGCGG
S. griseus	Ser	TGA	GGAGGGT	TG	CCCG	AGC GGCCTA	AGGG	A	ACGGT	CT TGA AA	ACCGT
S. lividans	Tyr	GTA	GGCGGTG	TG	CCCG	AGC GGCCAA	AGGG	A	GCAGA	CT GTA AA	TCTGC
S. coelicolor	*glnT α*	CTG	TGGCCTA	TG	GTGT	AAT GGC A	GCAC	G	ACTGA	TT CTG GT	TCAGT
S. coelicolor	*glnT β*	CTG	TGGTCTA	TG	GTGT	AATT GGC A	GCAC	G	ACTGA	TT CTG GT	TCAGT
S. lividans	*glnT α*	CTG	TGGCCTA	TG	GTGT	AATT GGC A	GCAC	G	ACTGA	TT CTG GT	TCAGT
S. lividans	*glnT β*	CTG	TGGTCTA	TG	GTGT	AATT GGC A	GCAC	G	ACTGA	TT CTG GT	TCAGT
S. rimosus	*glnT α*	CTG	TGGGCTA	TG	GTGT	AATT GGC A	GCAC	G	ACTGA	TT CTG GT	TCAGT
S. rimosus	*glnT β*	CTG	TGGGCTA	TG	GTGT	AATT GGC A	GCAC	G	AGTGA	TT CTG GT	TCATT
S. coelicolor	*gluT α,β,γ*	CTC	GCCCCCG	TT	GTGT	AGC GGCCTA	GCAC	G	CCGCC	CT CTC AA	GGCGG
S. lividans	*gluTa,b,g*	CTC	GCCCCCG	TT	GTGT	AGC GGCCTA	GCAC	G	CCGCC	CT CTC AA	GGCGG
S. rimosus	*gluT α,β,γ*	CTC	GCCCCCG	TT	GTGT	AGC GGCCTA	GCAC	G	CCGCC	CT CTC AA	GGCGG
S. coelicolor	Gly	CCC	GCGGGTG	TA	GTTC	AAT GGT A	GAAC	A	TCAGC	TT CCC AA	GCTGA
S. lividans	*glyU α,β*	GCC	GCGGACG	TA	GCTC	A CTTGGT A	GAGC	G	CAACC	TT GCC AA	GGTTG
S. lividans	*glyT*	TCC	GCGTTGG	TG	GTCC	AA GG AA	AGAC	G	CCCCA	CT TCC CG	TGGGG
S. coelicolor	Thr	GGT	GCCCCTA	TA	GCTC	AGTC GGT A	GAGC	G	TCTCC	AT GGT AA	GGAGA
phi-C31	Thr	CGT	GCCTCCC	TA	GCTC	AGTTCGGT TA	GAGC	G	CCTGT	TT CGT AA	TCAGG
S. livida	*cysT*	GCA	GGTGGAG	TG	GCCG	AG AGGC G	AGGC	A	ACGGC	CT GCA AA	GCCGT
S. lividans	*valT α,γ*	GAC	GGACGAT	TA	GCTC	AGC GG GA	GAGC	G	CTTCC	CT GAC AC	GGAAG
S. lividans	*valT β*	GAC	GCGCGAT	TA	GCTC	AGC GG GA	GAGC	G	CTTCC	CT GAC AC	GGAAG
S. lividans	Asp	GTC	GGTCCTG	TG	GAGC	AGTTTGGAGT	GCTC	G	CCACC	CT CTC AA	GGTGG
S. lividans	*lysT α,β*	CTT	GCGCCGC	TA	GCTC	AGTT GGT TA	GAGC	A	GCTGA	CT CTT AA	TCAGC
S. lividans	*lysT γ*	CTT	GCGCCGC	TA	GCTC	AGTT GGT TA	GAGC	A	GCTGA	CT CTT AA	TCAGC
S. lividans	*asnT α,β*	GTT	TCCTCGG	TA	GCTC	AATT GGC A	GAGC	A	GCCGG	CT GTT AA	CCGGC
S. lividans	Arg	ACG	GCACTCG	TA	GCTT	AAC GG ATA	GAGC	A	TCTGA	CT ACG GA	TCAGA
S. venezuelae	Arg	CCT	GCCTTCG	TA	GCTC	AG GGGT A	GAGC	A	CCGCT	CT CCT AA	AGCGG
S. venezuelae	Arg	CCT	GCCTTCG	TA	GCTC	AG GGG ATA	GAGC	A	CCGCT	CT CCT AA	AGCGG
S. rimosus	Arg	CCT	GCCTCCG	TA	GCTC	AG GGG ATA	GAGC	A	CCGCT	CT CCT AA	AGCGG
S. griseus	Trp	CCA	AGGGTCG	TA	GCTC	AATT GGT A	GAGC	A	CTGGT	CT CCA AA	ACCAG
S. galbus	Trp	CCA	AGGGTCG	TA	GCTC	AATT GGT A	GAGC	A	CCGGT	CT CCA AA	ACCGG
S. griseus	Met	CAT	CGCGGGG	TG	GAGC	AGCTCGGT A	GCTC	G	CTGGG	CT CAT AA	CCCAG
S. rimosus	Met	CAT	CGCGGGG	TG	GAGC	AGCTCGGT A	GCTC	G	CTGGG	CT CAT AA	CTCAG
S. pristinaespiralis	Pro	CGG	CGGGGTG	TC	GCGC	AGCTTGGC A	GCGC	G	CTTCG	TT CGG GA	CGAAG
S. ambofaciens	Pro	CGG	CGGGGTG	TG	GCGC	AGCTTGGT A	GCGC	G	CTTCG	TT CGG GA	CGAAG
Invariant and Semi-invariant nucleotides[1]				T-	GY-	ARY GGY A	R RC	R		YT R	

[1] A dash (-) indicates bases that are invariant or semi-invariant in some organisms but are not considered invariant or semi-invariant in *Streptomyces*

[2] Alignment is centered around the invariant G18-G19 nucleotides, and adjusted to maximize the alignment of additional invariant and semi-invariant bases in this region

[3] If sequence has not been deposited in Genbank, the relevant reference is indicated

Variable stem-loop	T stem-loop			Aminoacyl stem	3'end	tRNA	Genbank Accession #³
44-48	49-53	54-60	61-65	66-72	73-76		
TGGCCTTCATGGCCG	T GCCGG	TTCGAGT	CCGGC	TCCGGGC	AC	Leu	This work
TGCCCCTTCGAGGGCG	T GCCGG	TTCAAGT	CCGGC	TCCGGGC	ACCA	Leu	Y00209
TGCCCCTTCGAGGGCG	T GCCGG	TTCAAGT	CCGGC	TCCGGGC	ACCA	Leu	Ueda *et al.*, 1992
TGCCCCTCGGGGGCG	T GCCGG	TTCGAGT	CCGGC	TCCGGGC	AC	Leu	M80629
TGCCCTTTATCGGGCG	T GGGGG	TTCAAGT	CCCCC	CTCGGAC	ACCA	leuUα	This work (U59627)
TGCCCTTTATCGGGCG	T GGGGG	TTCAAGT	CCCCC	CTCGGAC	AC	leuUβ	AL031013
TTTGGGTCTTAAAGCCCATC	GAGGG	TTCAAAT	CCCTC	CGCCTCC	GC	Ser	X70689
CGTGGTGGCGACATCAC	C GTGGG	TTCAAAT	CCCAC	ACCCTCC	GC	Ser	M86368
CGGCTCAGCCTT	C CCAGG	TTCGAAT	CCTGG	CGCCGCC	AC	Tyr	X68439
TAG	T CTAGG	TTCGAGT	CCTGG	TAGGCCA	GC	*glnT* α	AL031124
TAG	T CTAGG	TTCGAGT	CCTGG	TAGACCA	GC	*glnT* β	AL031124
TAG	T CTAGG	TTCGAGT	CCTGG	TAGGCCA	GC	*glnT* α	X58873
TAG	T CTAGG	TTCGAGT	CCTGG	TAGACCA	GC	*glnT* β	X58873
TAG	T CTAGG	TTCGAGT	CCTGG	TAGCCCA	GC	*glnT* α	X53649
TAG	T CTAGG	TTCGAGT	CCTGG	TAGCCCA	GC	*glnT* β	X53649
TAG	C GCCGG	TTCGAAT	CCGGT	CGGGGGT	AC	*gluT* α,β,γ	AL031124
TAG	C GCCGG	TTCGAAT	CCGGT	CGGGGGT	AC	*gluTa,b,g*	X58873
TAG	C GCCGG	TTCGAAT	CCGGT	CGGGGGT	AC	*gluT* α,β,γ	X53649
TAG	C GCGAG	TTCGATT	CTCGT	CACCCGC	TCCA	Gly	X51702
AGGT	C GCGAG	TTCGAGC	CTCGT	CGTCCGC	TC	*glyU* α,β	X52072, X658875
AAA	T GCAGG	TGCAAGG	CCTGC	CCAGCGC	TC	*glyT*	X52071
AGGT	C AACGG	TTCGATT	CCGTT	TGGGGGC	TCCA	Thr	AI031031
GGGT	C GGCGG	TTCGAAT	CCGTC	GGGGGGC	TC	Thr	X91149
CTA	C ACGGG	TTCAAAT	CCCGT	CTCCACC	TCCA	*cysT*	X52072
AGGT	C ACTGG	TTCAATC	CCAGT	ATCGTCC	AC	*valT* α,γ	X52072
AGGT	C ACTGG	TTCAATC	CCAGT	ATCGCGC	AC	*valT* β	X52072
AGGC	C GCGGG	TTCAAAT	CCCGT	CAGGACC	GC	Asp	X58875
GGGT	C CGGGG	TTCGAGT	CCCTG	GCGGCGC	AC	*lysT* α,β	X52073-4
GGGT	C CGGGG	TTCGAGT	CCCTG	GCGGCGC	ACC	*lysT* γ	X63142
AGGT	T ACTGG	TTCGAGT	CCAGT	CCGGGGA	GC	*asnT* α,β	X52070
AGGT	T GCAGG	TTCGAAT	CCTGC	CGAGTGC	AC	Arg	X70689
GTGT	C GCAGG	TTCGAAT	CCTGC	CGGGGGC	ACCA	Arg	AJ000048
GTGT	C GCAGG	TTCGAAT	CCTGC	CGGGGGC	AC	Arg	AJ000050
GTGT	C GCAGG	TTCGAAT	CCTGC	CGGGGGC	ACCA	Arg	X67954
CGGT	T GGGGG	TTCAAGT	CCCTC	CGGCCCT	GC	Trp	X72787
CGGT	T GGGGG	TTCAAGT	CCCTC	CGGCCCT	GC	Trp	X95916
AGGT	C GCAGG	TTCAAAT	CCTGT	CCCCGCT	ACCA	Met	X04543
AGGT	C GCAGG	TTCAAAT	CCTGT	CCCCGCT	AC	Met	M32254
AGGT	C GTGGG	TTCAAAT	CCCGC	CACCCCG	AC	Pro	YO9425
AGGT	C GTGGG	TTCAAAT	CCCGC	CACCCCG	AC	Pro	M22964
	Y RG	TTCRA	Y CY		C--		

Table III.4.2 Invariant and Semi-Invariant Bases in *Streptomyces* tDNAs

Invariant/Semi-invariant base	Exceptions (base specified, tRNA)
8T	None
10G	C in *S. griseus* Ser(TGA) and *S. lividans* Tyr (GTA)
11Y	A in *S. lividans* Asp(GTC), *S. rimosus* and *S. griseus* Met
14A	None
15R	missing in *S. lividans glyU* α,β
16Y	missing in *S. lividans glyT, cysT; S. venezuelae* and *S. rimosus* Arg(CCT)
18G	missing in all *bldA* tRNAs (Leu(TAA))
19G	None
20Y	missing in several tRNAs, A in *S. lividans* Asp(GTC)
21A	missing in several tRNAs
22R	G>>A, T in *S. lividans* Ser(GCT)
24R	T in *S. lividans* Asp(GTC), *S.rimosus* and *S. griseus* Met
25C	G in *S. griseus* Ser(TGA) and *S. lividans* Tyr(GTA)
26R	None
32Y	A in *S. coelicolor* Thr(GGT)
33T	None
37R	C in *S. lividans glyT*
48Y	None
52R	None G>>A
53G	None
54T	None
55T	None
56C	None
57R	None
58A	None
60Y	T>>C, G in *S. lividans glyT*
61C	None
62Y	None C>>T
74C	None

Bases excluded	Reason
9R	majority contain R (39), remaining 11 contain Y
12Y	22/50 contain Y (invariant/semi-invariant only in *Saccharomyces cerevisiae* mitochondrial tRNAs)
75C	40/50 tRNAs missing this base
76A	41/50 tRNAs missing this base

sequence of the 50 *Streptomyces* tRNA genes the bases considered invariant and semi-invariant are; 8T, 10G, 11Y, 14A, 15R, 16Y, 18G, 19G, 20Y, 21A, 22R, 24R, 25C, 26R, 32Y, 33T, 37R, 48Y, 52R, 53G, 54T, 55T, 56C, 57R, 58A, 60Y, 61C, 62Y, and 74C (Table III.4.2), with an R indicating a purine (A or G) and a Y indicating a pyrimidine (C or T). Of these bases, 10G, 11Y, 20Y, 22R, 25C, 32Y, 37R, 60Y, and 62Y have up to three tDNAs that differ at any one of these positions. The D-loop of the *Streptomyces* tRNAs has between 11 *(S. lividans* $\text{tRNA}^{\text{Ser}}_{\text{GCT}}$) and 6 *(S. lividans glyT)* bases, so many of the invariant and semi-invariant bases in this region (15R, 16Y, 18G, 20Y, and 21A) are missing in numerous tRNAs. However, when present, the bases typically conform to the expected pattern. The base 9R, which is invariant or semi-invariant in most organisms, is not conserved in *Streptomyces* as 11/50 tDNAs contain a pyrimidine in this position. However, this base is also not conserved in *E. coli* tRNAs, so lack of conservation at this position is not without precedent. The absence of the C75-A76 is unusual as these bases are typically conserved in other prokaryotes, but since only 10 of the 50 *Streptomyces* tDNAs contain the complete CCA terminus it is apparently not unusual for *Streptomyces* tRNA genes.

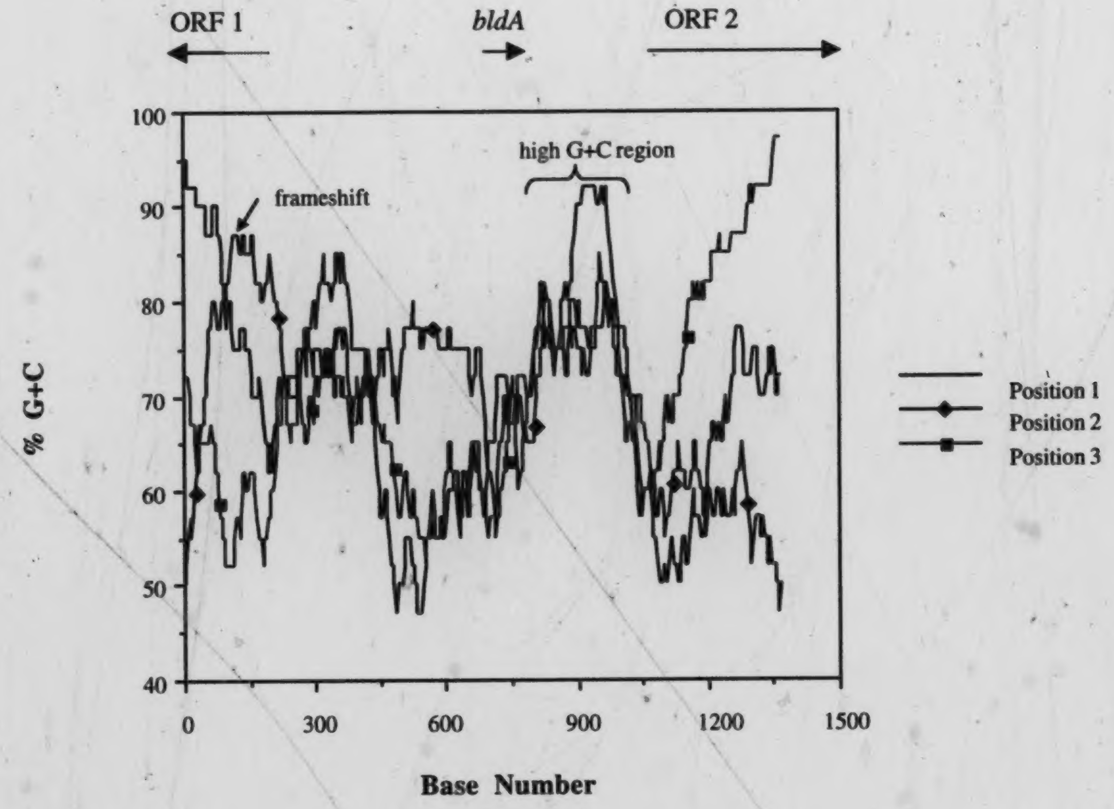
The *bldA* tDNAs from *S. clavuligerus, S. coelicolor, S. lividans* and *S. griseus* all are missing the invariant and semi-invariant bases 18G and 21A. Other *Streptomyces* tRNA genes are missing the 21A so its absence is apparently not uncommon, however, these are the only tRNA genes that are missing the highly conserved 18G. The D-loop of these tRNAs is only 7 bases while the majority of tRNAs contain D-loops of 8-9 bases, however, there are several other tRNAs with D-loops of only 7 bases and one with only 6 bases *(S. lividans glyT)*, so the shortened D-loop doesn't appear to be unusual. All of the other invariant and semi-invariant bases identified in *Streptomyces* are conserved in the *bldA* tRNA genes.

A search for open reading frames using the Frame analysis program designed by Bibb *et al.* (1984) indicated the presence of two partial open reading frames on the 1.5 kb *Sal*I fragment (Figure III.4.3). Frame analysis is based on the observation that *Streptomyces* open reading frames have a G+C distribution of ~70% in the first codon position, ~50% G+C in the second codon position and ~90% G+C in the third codon position. If a *Streptomyces* DNA fragment is graphed by G+C content against triplet position the relative location and orientation of potential reading frames can be easily identified as they show a 70%, 50%, 90% G+C distribution. For the 1.5 kb fragment there are two potential divergent open reading frames at either end of the DNA fragment.

Initial Frame analysis of the first putative open reading frame, which is upstream and divergent to the *bldA* tRNA gene, indicated that the open reading frame started around base 250 and extended through to base 1 with a frameshift of +1 or -2 within the open reading frame. The frameshift is predicted as the open reading frame initially starts in the 1-3-2 frame and shifts to the 3-2-1 frame. There is a potential ATG start codon at base 202 that is in the 1-3-2 frame and is located near where the open reading frame is predicted to begin. Additional sequencing of both strands to identify the frameshift was carried out using a denaturing gel containing 40% formamide as well as resequencing by a sequencing service. Resequencing of this region failed to identify any sequencing errors so it is unlikely that the open reading frame starts in the 1-3-2 frame and shifts to the 3-2-1 frame. Alternate GTG and ATG start codons are present further downstream at nucleotide 120 and 111 respectively. Both of these start codons are downstream from the

Figure III.4.3

Frame analysis of the 1.5 kb *SalI* fragment containing the *bldA* gene. The plain black line represents the first position of each triplet, the blue (♦) line the second position and the red (■)line the third position. The *bldA* tRNA gene and two potential open reading frames are indicated by arrows. The direction of the arrows indicates the direction of transcription. The potential +1 or –2 frameshift in ORF1 is indicated. The region that has an overall G+C content of ~80% is also indicated.


ORF 1
bldA
ORF 2
high G+C region
frameshift
% G+C
100
90
80
70
60
50
40
0
300
600
900
1200
1500
Base Number
Position 1
Position 2
Position 3

predicted frameshift and are in the 3-2-1 frame. The corresponding amino acid sequence of the ORF starting at GTG120 shows significant homology to ORFJI1 from *S. coelicolor* (Lawlor, 1987) (Figure III.4.4). ORFJI1 is located upstream from and divergent to *bldA*, similar to the position and orientation of ORF1 in relation to *bldA* in *S. clavuligerus*. In *S. coelicolor,* ORFJI1 was predicted to consist of two peptide sequences, the first open reading frame encodes a small 25 amino acid peptide which is separated from the larger downstream open reading frame by two in-frame stop codons. In contrast, the *S. clavuligerus* homologue, consists of a single open reading frame and encodes the amino acids Glu and Gln instead of stop codons. While it is possible that the reported *S. coelicolor* sequence is accurate and the twenty five amino acid peptide is no longer part of the larger open reading frame it also seemed likely that the reported sequence contains errors as the stop codons differ in each case from the codons for Glu and Gln in the *S. clavuligerus* ORF1 by a single nucleotide. In order to verify the sequence of ORFJI1, a 528 bp fragment was amplified from *S. coelicolor* chromosomal DNA by PCR and the purified product was sequenced. Several sequencing errors were identified including a +1 and –1 frameshift within the same codon as well as errors that specified the two stop codons. The corrected sequence for ORFJI1 is shown in Figure III.4.5. Because of the high degree of homology between ORFJI1 from *S. coelicolor* and ORF1 from *S. clavuligerus* it seemed likely that the primer that is homologous to the nucleotide sequence of codons 92-97 of ORFJI1 (BKL85, Figure III.4.5) would also hybridize to same region of ORF1, and could be used to obtain additional sequence of ORF1 past the *Sal*I restriction site. The homologous region of ORF1 was amplified by PCR using the primer BKL85 which was designed from the *S. coelicolor* sequence along with the primer

Figure III.4.4

Comparison of the nucleotide and amino acid sequence of ORF1 *(S. clavuligerus)* and ORFJI1 *(S. coelicolor)*. The sense strand is given with the amino acid sequence above *(S. clavuligerus)* and below *(S. coelicolor)*. The differences in encoded amino acids are indicated in grey. The putative ATP binding motif of ORFJI1 is underlined.

```
                 fM  T   T   M   P   I   A   H   R   A   T   A   V   A   A
S. clavuligerus  GTG ACC ACC ATG CCC ATC GCA CAC CGC GCC ACC GCC GTG GCC GCC
S. coelicolor    GTG ACC ACC GCA CCC ATC GCC GAC CGG TCC ACC CTC GTG GCC GCG
                 fM  T   T   A   P   I   A   D   R   S   T   L   V   A   A

                 R   A   T   E   L   S   K   V   Y   G   E   G   E   T   Q
S. clavuligerus  CGC GCC ACG GAA CTG TCC AAG GTG TAC GGG GAG GGC GAG ACC CAG
S. coelicolor    CGC GCC ACG GAG CTT TCC AAG ATC TAC GGC TAG GGC GAG ACC TAG
                 R   A   T   E   L   S   K   I   Y   G   *   G   E   T   *

                 V   V   A   L   D   R   V   S   V   D
S. clavuligerus  GTG GTG GCC CTG GAC CGG GTC AGC GTC GAC
S. coelicolor    GTG GTC GCC TTG GAC CGG GTC TCC ATC GAC TTC CGG CAG GCC GAG
                 M   V   A   L   D   R   V   S   I   D   F   R   Q   A   E

S. coelicolor    CTC ACC GCG ATC ATG GGC CCC TCC GGC TCC GGC AAG TCC ACG CTG
                 L   T   A   I   M   G   P   S   G   S   G   K   S   T   L

S. coelicolor    ATG CAC TGC GTC GCG GGC TGG GAC ACC TTC TGG GCC GGC TCG GTG
                 M   H   C   V   A   G   W   D   T   F   S   A   G   S   V

S. coelicolor    CGC ATC GGC GAG ACC GAG CTG GGC TCG CGC AAG GAC AAG CAG CTC
                 R   I   G   E   T   E   L   G   S   R   K   D   K   Q   L

S. coelicolor    ACC AAG CTG CGC CGG GAC AAG ACG GCT CAC
                 T   K   L   R   R   D   K   T   A   H
```

Figure III.4.5

Corrected sequence of ORFJI1 from *S. coelicolor*. The differences in nucleotide and amino acid sequence from previously published sequence (Lawlor, 1987) are indicated in bold. The primers BKL84 (5'-CGGCCGTCATCCGCAGGT-3') and BKL85 (5'-CTTGTCCCGGCGCAGCTT-3') were designed from previously published sequence, and are indicated by arrows. Regions where reliable sequence was not obtained are indicated in grey. The putative ATP binding motif previously identified is indicated by underlined amino acids.

BKL84
⟶
CGGCCGTCATCCGCAGGTATGACACCGGCCCGGCCCGTCCGACCGAAGTCGCCCCCCTGA
GCCGGCAGTAGGCGTCCATACTGTGGCCGGGCCGGGCAGGCTGGCTTCAGCGGGGGGACT

ACCGGAACAGCGGCTGACTACAGGGCGCCGTTCGACCGGGACGATTGACAACGTCCCCGA
TGGCCTTGTCGCCGACTGATGTCCCGCGGCAAGCTGGCCCTGCTAACTGTTGCAGGGGCT

GCACGCACCGTGCCGACCGCACGCACCGTGCCCAGCACGCACCGTGCCCAGCACACACCG
CGTGCGTGGCACGGCTGGCGTGCGTGGCACGGGTCGTGCGTGGCACGGGTCGTGTGTGGC

ORFJI1 fM
TGCCACCGCACGCACCGTGCCGACCGCACACGCCGTACCGACAGGAGCGCCCTTCCCGTG
ACGGTGGCGTGCGTGGCACGGCTGGCGTGTGCGGCATGGCTGTCCTCGCGGGAAGGGCAC

T T A P I A D R S T L V A A R A T E L S
ACCACCGCACCCATCGCCGACCGGTCCACCCTCGTGGCCGCGCGCCACGGAGCTTTCC
TGGTGGCGTGGGTAGCGGCTGGCCAGGTGGGAGCACCGGCGCGCGGTGCCTCGAAAGG

K I Y G **Q** G E T **Q** **V** V A L D R V S I D F
AAGATCTACGGC**C**AGGGCGAGACC**C**AGGTGGTCGCC**C**TGGACCGGGTCTCCATCGACTTC
TTCTAGATGCCG**G**TCCCGCTCTGG**G**TCCACCAGCGG**G**ACCTGGCCCAGAGGTAGCTGAAG

R Q A E L T A I M G P S G S G K S T L M
CGGCAGGCCGAGCTCACCGCGATCATGGGCCCCTCCGGCTCCGGCAAGTCCACGCTGATG
GCCGTCCGGCTCGAGTGGCGCTAGTACCCGGGGAGGCCGAGGCCGTTCAGGTGCGACTAC

H C V A G **L** D T F S **S** G S V R I G E T E
CACTGCGTCGCGGGC**CTG**GACACCTTCT**C**G**T**CCGGCTCGGTGCGCATCGGCGAGACCGAG
GTGACGCAGCGCCCG**GAC**CTGTGGAAGA**G**C**A**GGCCGAGCCACGCGTAGCCGCTCTGGCTC

L G S **L** K D K Q L T K L R R D K
CTGGGCTCGC**T**CAAGGACAAGCAGCTCACCAAGCTGCGCCGGGACAAG
GACCCGAGCG**A**GTTCCTGTTCGTCGAGTGGTTCGACGCGGCCCTGTTC
⟵
BKL85

NTR21 which is homologous to the region 226-246 bases upstream from the ORF1 start codon. The amplified fragment was purified and sequenced. The additional sequence obtained, along with corresponding amino acid sequence is shown in Figure III.4.6, and an amino acid comparison of ORFJI1 and ORF1 is shown in Figure III. 4.7. The two amino acid sequences share 94.5% similarity and 90.0% identity with no gaps. Although no potential function can be assigned to either ORF they both contain a putative ATP binding motif.

The second partial open reading frame is located downstream from and in the same orientation as the *bldA* tRNA gene. The partial amino acid sequence of this gene shows 82.5% similarity and 81.4% identity with 3 gaps (gap creation penalty of 12 and gap extension penalty of 4) with the gene located downstream from the *S. griseus bldA* tDNA (Figure III.4.8) (Kwak *et al.*, 1996). Although ORF2 shows no significant homology to any other protein or nucleotide sequence in the databases, this protein may be an integral membrane protein as the *S. griseus* homologue has been completely sequenced and reportedly contains seven closely spaced hydrophobic domains. Consistent with this prediction, a search for motifs with the GCG Wisconsin program indicated that the *S. griseus* ORF contains a prokaryotic membrane lipoprotein lipid attachment site. Based on the apparent homology and arrangement of genes in the vicinity of the *bldA* gene from several *Streptomyces* species, it appears that this region has been highly conserved among streptomycetes.

One other region of note in the *S. clavuligerus bldA* cloned DNA is a segment of DNA downstream of the *bldA* gene that is extremely high (~80%) in G+C content (See Figure III.4.3), and contains several large, 13-20 base imperfect direct repeats

Figure III.4.6

Additional sequence of ORF1 from *S. clavuligerus*. The sequence was obtained by PCR amplification of *S. clavuligerus* chromosomal DNA using the primers NTR21 (5'-ACACCGCAGCGAAGAGTGA-3'), which is homologous to a region upstream of ORF1, and BKL85 (5'-CTTGTCCCGGCGCAGCTT-3') which is complementary to codons 92-97 of *S. coelicolor* ORFJI1. The PCR product was sequenced by a sequencing service using the same primers. The sequence from NTR21 to the first putative start codon of ORF1 is not shown as it is identical to the sequence previously reported (Figure III.4.1). The location of the primer BKL85 binding site is indicated in grey. It corresponds to the homologous region of *S. coelicolor* ORFJI1 and may differ in *S. clavuligerus*. The sequence complementary to the primer, as well as a few nucleotides 3' of the primer are also colored grey as this sequence was considered unreliable. The putative ATP binding motif is indicated by underlined amino acids and the *Sal*I restriction site which indicates the end of the *bldA* cloned fragment is also underlined.

```
fM T  T  M  P  I  A  H  R  A  T  A  V  A  A  R  A  T  E  L
GTGACCACCATGCCCATCGCACACCGCGCCACCGCCGTGGCCGCCCGCGCCACGGAACTG
CACTGGTGGTACGGGTAGCGTGTGGCGCGGTGGCGGCACCGGCGGGCGCGGTGCCTTGAC

S  K  V  Y  G  E  G  E  T  Q  V  V  A  L  D  R  V  S  V  D
TCCAAGGTGTACGGGGAGGGCGAGACCCAGGTGGTGGCCCTGGACCGGGTCAGCGTCGAC
AGGTTCCACATGCCCCTCCCGCTCTGGGTCCACCACCGGGACCTGGCCCAGTCGCAGCTG
                                                       SalI

F  R  Q  A  E  F  T  A  I  M  G  P  S  G  S  G  K  S  T  L
TTCCGGCAGGCCGAGTTCACCGCGATCATGGGGCCCTCGGGCTCCGGCAAGTCGACCCTG
AAGGCCGTCCGGCTCAAGTGGCGCTAGTACCCCGGGAGCCCGAGGCCGTTCAGCTGGGAC

M  H  C  V  A  G  L  D  S  F  S  S  G  S  V  R  I  G  E  T
ATGCACTGCGTCGCCGGGCTCGACAGCTTCAGCTCGGGCTCCGTCCGGATCGGCGAGACC
TACGTGACGCAGCGGCCCGAGCTGTCGAAGTCGAGCCCGAGGCAGGCCTAGCCGCTCTGG

E  L  G  S  L  K  D  K  Q  L
GAGCTGGGCTCCCTCAAGGACAAGCAGCTCACCAAGCTGCGCCGGGACAAG
CTCGACCCGAGGGAGTTCCTGTTCGTCGAGTGGTTCGACGCGGCCCTGTTC
                                ◄──────────────────
                                      BKL85
```

Figure III.4.7

Alignment of the amino acid sequence of ORF1 from *S. clavuligerus* and ORFJI1 from *S. coelicolor.* Identical amino acid residues are indicated by |, highly similar residues are indicated by :, and similar residues are indicated by ·. The putative ATP binding motifs are underlined.

```
S. clavuligerus ORF1     MTTMPIAHRATAVAARATELSKVYGEGETQ     30
                         ||| ||| |·| ||||||||||:||:||||
S. coelicolor ORFJI1     MTTAPIADRSTLVAARATELSKIYGQGETQ     30


S. clavuligerus ORF1     VVALDRVSVDFRQAEFTAIMGPSGSGKSTL     60
                         ||||||||:||||||.||||||||||||||
S. coelicolor ORFJI1     VVALDRVSIDFRQAELTAIMGPSGSGKSTL     60


S. clavuligerus ORF1     MHCVAGLDSFSSGSVRIGETELGSLKDKQL     90
                         ||||||||·|||||||||||||||||||||
S. coelicolor ORFJI1     MHCVAGLDTFSSGSVRIGETELGSLKDKQL     90
```

Figure III.4.8

Alignment of the amino acid sequence of ORF2 from *S. clavuligerus* and an ORF located downstream of the *S. griseus bldA* gene. Identical amino acid residues are indicated by |, highly similar residues are indicated by :, and similar residues are indicated by ·.

```
S. clavuligerus ORF2   MRSSNPVFSRRGFSRDNGIAGFNAQQQAGSPYAGN   35
                       |||||||||||||||||| |||||· |||·| ·||
S. griseus ORF         MRSSNPVFSRRGFSRDNGHAGFNATPQAGAPATGN   35

S. clavuligerus ORF2   PYAQG·AAN[illegible]ATNPYAPADTQLGAAQAPAHR·VM   68
                       ||||| |||||||||||··|·| ||:||||:· ||
S. griseus ORF         PYAQGTAANPYATNPYAQTDIQ·GAPQAPARPDVM   69

S. clavuligerus ORF2   TIDDVVARTAMTLGTVVVTAMLAWFLLPVD   98
                       ||||||·|||||||||:|·|  ||:|||||
S. griseus ORF         TIDDVVTRTAMTLGTVIVAATAAWWLLPVD   99
```

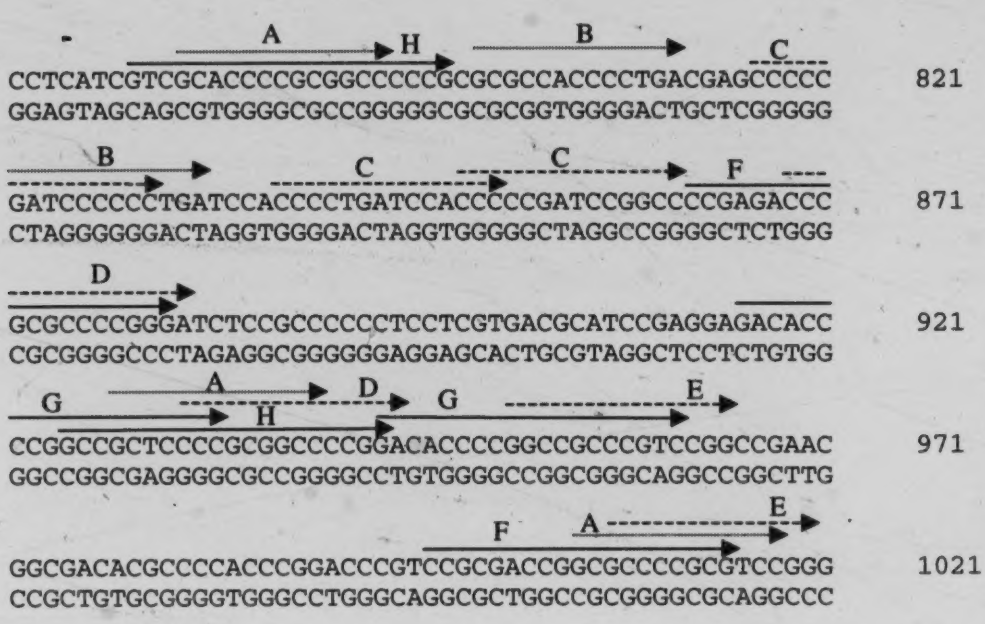
(Figure III.4.9). The function of this region is unknown but it is worthy to note that one strand contains 53.5% G bases which often occur in runs of three or four nucleotides. Guanosine residues can potentially form non-Watson Crick base pairs with most bases. The fact that the template strand containing predominately G residues could not be sequenced by conventional techniques indicates that secondary structures can form under some conditions. Although this region doesn't show end-to-end homology to the corresponding region from *S. coelicolor* or *S. griseus*, there are small regions of homology which indicate that some of the repetitive sequences observed in the *S. clavuligerus* sequence are somewhat conserved in the three species.

III.5 GENERATION OF A *bldA* NULL MUTANT STRAIN

In order to determine the function of the *bldA* tRNA in *S. clavuligerus*, a null mutant strain was generated by gene replacement. The general scheme for generating the recombinant plasmid to be used for gene replacement is shown in Figure III.5.1. The upstream flanking sequence of the *bldA* gene (left arm) was purified as a *Sal*I- *Nco*I 540 bp fragment from the plasmid p9S- by the crush and soak procedure. The fragment containing sequence downstream from the *bldA* gene was generated by PCR using the plasmid p9S+ as template DNA. The primers used to amplify the downstream sequence (right arm) were the universal sequencing primer (5'-GTAAAACGACGGCCAGT-3') and NTR5 (5'-GCGC**GGTACC**GCACTCTCCGTAACGAGA-3'), which is complementary to the sequence immediately downstream of the *bldA* tRNA and contains a 5' non-homologous tail (underlined) containing a *Kpn*I restriction site (bold). The PCR product was digested with the restriction enzymes *Bam*HI and *Kpn*I to generate the 3' flanking fragment. The replacement fragment, containing the gene for apramycin resistance was purified from

Figure III.4.9

Direct repeats in G+C rich region from base 772-1021 of the 1.5 kb *Sal*I *bldA*-containing cloned DNA. Repeats of 19-20 bases contain 2-3 mismatches and repeats of 13-14 bases contain 2 mismatches. The two 13 base repeats are indicated by arrowheads with dotted lines (A and B). The three 14 base repeats are indicated by arrowheads with dashed lines (C, D, and E). The two 19 base repeats are indicated by arrowheads with solid lines (F and G), and the single 20 base repeat is indicated by a arrowhead with solid line (H).



CCTCATCGTCGCACCCCGCGGCCCCCGCGCGCCACCCCTGACGAGCCCCC 821
GGAGTAGCAGCGTGGGGCGCCGGGGGCGCGCGGTGGGGACTGCTCGGGGG

GATCCCCCCTGATCCACCCCTGATCCACCCCCCGATCCGGCCCCGAGACCC 871
CTAGGGGGGACTAGGTGGGGACTAGGTGGGGGCTAGGCCGGGGCTCTGGG

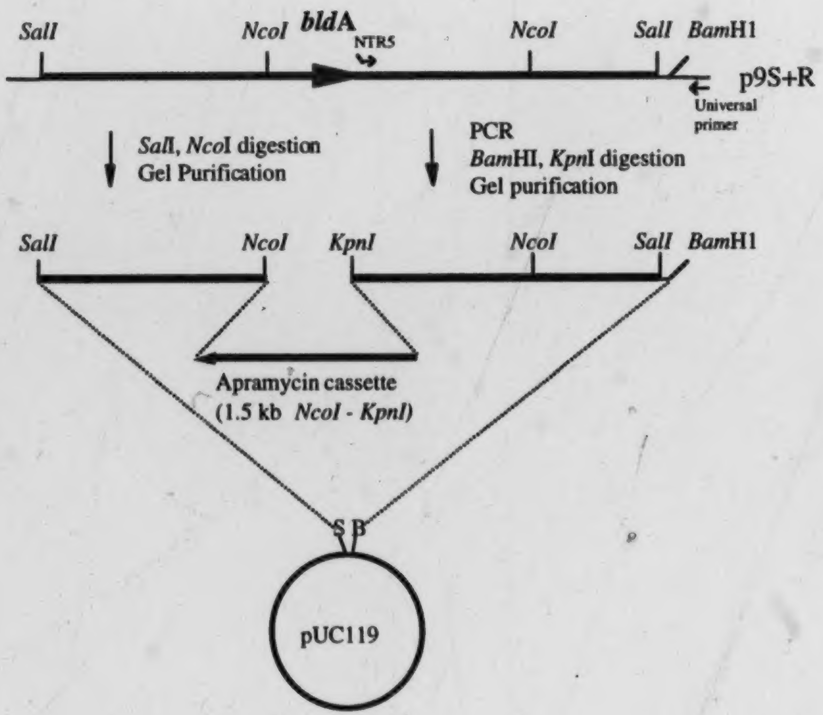
GCGCCCCGGGATCTCCGCCCCCCTCCTCGTGACGCATCCGAGGAGACACC 921
CGCGGGGCCCTAGAGGCGGGGGGAGGAGCACTGCGTAGGCTCCTCTGTGG

CCGGCCGCTCCCCGCGGCCCCGGACACCCCGGCCGCCCGTCCGGCCGAAC 971
GGCCGGCGAGGGGCGCCGGGGCCTGTGGGGCCGGCGGGCAGGCCGGCTTG

GGCGACACGCCCCACCCGGACCCGTCCGCGACCGGCGCCCCGCGTCCGGG 1021
CCGCTGTGCGGGGTGGGCCTGGGCAGGCGCTGGCCGCGGGGCGCAGGCCC

Figure III.5.1

Subcloning strategy for the replacement of the *S. clavuligerus bldA* gene with the *apramycin* resistance gene cassette. The sequence upstream of the *bldA* gene was purified as a 540 bp *Sal*I-*Nco*I DNA fragment. The *apramycin* gene cassette was purified from pUC120Ap(Nco) as an *Nco*I-*Kpn*I fragment shown in grey with the arrow indicating the direction of transcription. The sequence downstream from the *bldA* gene was generated by PCR using the plasmid p9S+ as template DNA and the universal sequencing primer (5'-GTAAAACGACGGCCAGT-3') and NTR5 (5'-GCGC**GGTACC**GCACTCTCCGTAACGAGA-3') as primers. The DNA regions the primers are homologous to are indicated by arrows, with the direction of the arrows corresponds to the primer binding 5'⇨3'. The primer NTR5 contains the sequence for a *Kpn*I restriction site (this is represented as a non-homologous tail on the primer arrow and underlined on the primer sequence with the *Kpn*I site shown in bold). The PCR product was digested with *Kpn*I and *Bam*HI for subcloning. The sequences flanking the *bldA* gene and the *apramycin* gene cassette were ligated simultaneously into *Sal*I-*Bam*HI digested pUC119 to generate the recombinant plasmid for gene replacement.


SalI
NcoI
bldA
NTR5
NcoI
SalI
BamH1
p9S+R
Universal primer
SalI, NcoI digestion
Gel Purification
PCR
BamHI, KpnI digestion
Gel purification
SalI
NcoI
KpnI
NcoI
SalI
BamH1
Apramycin cassette
(1.5 kb NcoI - KpnI)
S B
pUC119

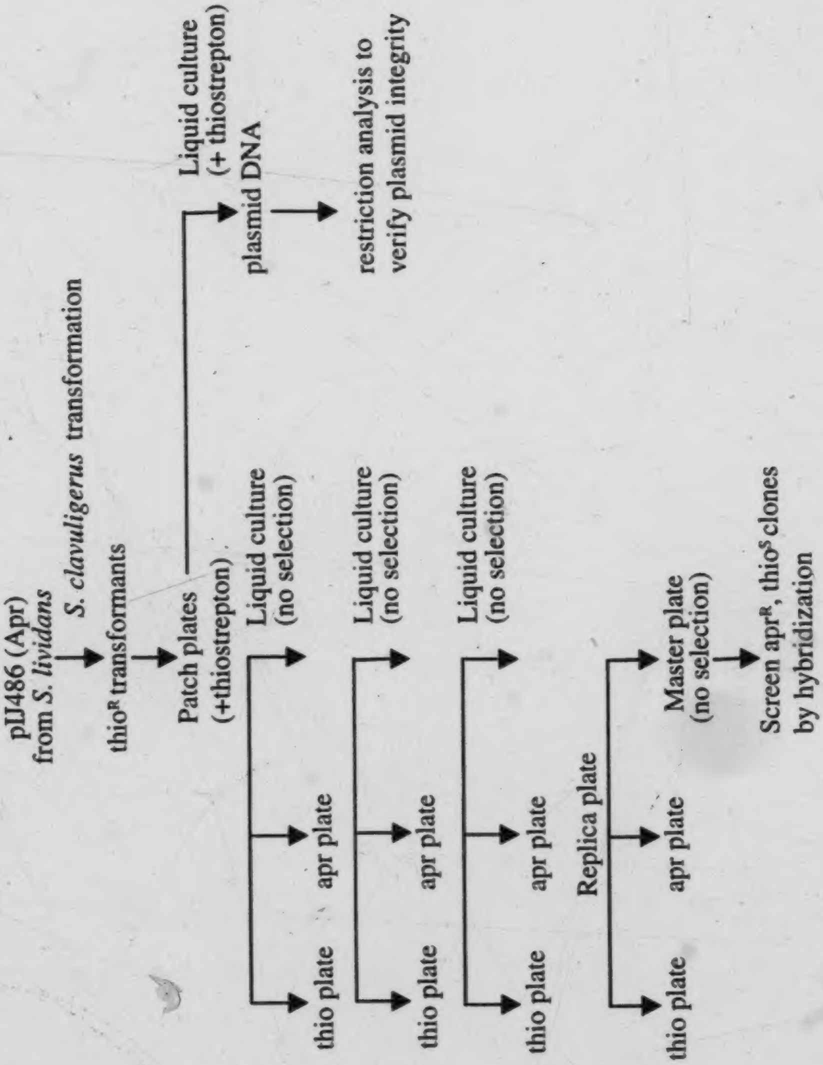
Subcloning Strategy for Gene Replacement pUC120Ap(Nco) as a *Kpn*I-*Nco*I fragment by the Gene Clean procedure. The three fragments (left arm, right arm, apramycin gene cassette) were ligated simultaneously into *Sal*I and *Bam*HI-digested pUC119. The ligation mix was transformed into *E. coli* MV1193 competent cells and plated on LB agar containing Xgal and IPTG for blue-white selection as well as apramycin for selection of insert-containing, plasmid transformants. There were only four white, apramycin resistant transformants, a number consistent with the expected low frequency of a 4-part ligation. The recombinant plasmids from the four transformants were isolated and checked by restriction digest analysis as well as by Southern hybridization (data not shown). Restriction analysis showed fragments of the expected size when digested with *Sal*I and *Sal*I-*Nco*I, and the appropriate fragments hybridized to probes specific to either the right or the left arm. Since the transformants were apramycin resistant and the 1.45 kb apramycin resistance gene was evident in the restriction digests, the recombinant plasmid digests were not probed with the apramycin gene cassette. Since all four recombinant plasmids from the apramycin resistant transformants showed the same restriction and hybridization patterns, one was chosen for further study and named pGD#1 for *G*ene *D*isruption *p*lasmid *#1*.

Two strategies were attempted to generate a *bldA* gene replacement strain. The first method was to use single-stranded DNA to transform *S. clavuligerus*. The reasoning was that the *bldA* gene in *S. coelicolor* had been successfully disrupted using single-stranded DNA (Leskiw *et al.*, 1993), and in other instances, single-stranded transformations had proved successful (Hillemann *et al.*, 1991). Also, the pGD#1 recombinant plasmid could be used directly without first subcloning the replacement fragment into a *Streptomyces* vector. In addition, *S. clavuligerus* is difficult to transform as it has several restriction systems and successful transformations require the passage of DNA first through *S. lividans*. Single-stranded DNA should not be restricted and is easily obtained from pUC119 vectors using the helper phage M13KO7. Since pGD#1 does not replicate in *Streptomyces,* it also functions as a suicide vector and any apramycin resistant transformants should result from a cross-over between the chromosome and homologous regions on the gene replacement plasmid. Single-stranded DNA was generated, the DNA was transformed into *S. clavuligerus* protoplasts, and apramycin was used to select for transformants. This procedure was attempted numerous times but no transformants were obtained. Upon testing, it was shown that the batch of *S. clavuligerus* protoplasts used for this procedure did not readily take up DNA, so it is unclear whether the absence of transformants was due to the protoplasts or because this method will not work in *S. clavuligerus*.

Rather than continue to pursue the single-stranded DNA strategy, a second strategy was used that involved subcloning of the gene replacement fragment into the *Streptomyces* vector pIJ486. Although pIJ486 is a high copy vector, it is commonly used for gene replacement in *S. clavuligerus* (Paradkar *et al.*, 1996; Paradkar and Jensen, 1995). *S. clavuligerus* is highly recombinogenic so that recombination between chromosomal DNA with homologous DNA on gene replacement plasmids will occur readily, making the use of a suicide vector unnecessary in this species. The 2.8 kb gene replacement fragment containing both the upstream and downstream *bldA* flanking sequences as well as the apramycin gene cassette was removed from pGD#1 as a *Hin*dIII-*Eco*RI fragment and ligated into similarly digested pIJ486. The ligation mixture was first transformed into *S. lividans* TK24, since *S. lividans* lacks restriction systems and will accept DNA originating from *E. coli.* Apramycin resistant transformants were selected and then patched onto R2YE plates containing apramycin and thiostrepton to ensure that the plasmid and insert were maintained and to generate spores. Spores from four transformants were cultured in liquid broth containing thiostrepton and the plasmids were isolated. A sample of purified plasmid DNA was digested with *Bgl*II, *Sal*I, and *Nco*I to verify that it contained insert of the appropriate size. pIJ486(Apr)#17 and pIJ486(Apr)#26 both contained the desired inserts and were used for subsequent transformations . Initial transformations into *S. clavuligerus* were performed using apramycin to select for the resulting transformants. While apramycin selection would select for transformants, it would not prevent plasmid loss if the apramycin gene cassette were to integrate into the host chromosome via double cross-over between the homologous sequences on the plasmid and *bldA* region on the chromosome. However, primary selection with apramycin yielded false positives, apparent *bld,* apramycin resistant colonies which proved to be apramycin sensitive upon further culturing. Subsequent attempts to introduce pIJ486(Apr) into *S. clavuligerus* were performed using thiostrepton to select for the transformation event. Primary transformants were subcultured onto MYM agar containing thiostrepton. The cultures from these master plates were used to inoculate thiostrepton-containing broth for plasmid isolation. The presence of plasmid containing insert was verified by restriction analysis using *Bgl*II and *Bgl*II-*Nco*I restriction enzymes. Several transformants that contained pIJ486(Apr) were then subcultured in liquid broth without selection to allow loss of the pIJ486 vector. The generalized scheme is shown in Figure III.5.2. Gene replacement by double crossing over with the homologous sequences in the *S. clavuligerus* chromosome, as well as

Figure III.5.2

Flow diagram outlining the procedure for generation of the *S. clavuligerus ΔbldA* mutant strain. The gene replacement plasmid pIJ486(Apr) was isolated from transformed *S. lividans* TK24 and transformed into *S. clavuligerus* protoplasts. Primary transformants were subcultured onto MYM agar containing thiostrepton. The cultures from these master patch plates were used to inoculate thiostrepton-containing broth for plasmid isolation. The presence of insert in these plasmids was verified by restriction analysis. Several transformants were subcultured in liquid broth without selection to allow loss of the pIJ486 vector. Gene replacement by double crossing over with the homologous sequences in the *S. clavuligerus* chromosome, as well as plasmid loss were monitored by plating aliquots of broth cultures on MYM agar plates containing apramycin or thiostrepton. When the sample plates indicated that the cultures were thiostrepton sensitive and apramycin resistant, aliquots were plated onto MYM agar without antibiotic. Isolated colonies were patched onto a master plate without selection and the master plates were used to replica plate MYM plates containing either thiostrepton or apramycin. Colonies displaying apramycin resistance and thiostrepton sensitivity were chosen for chromosomal DNA isolation and hybridization analysis to confirm the gene replacement.


pIJ486 (Apr)
from *S. lividans*
S. clavuligerus transformation
thio^R transformants
Patch plates
(+thiostrepton)
Liquid culture
(+ thiostrepton)
plasmid DNA
restriction analysis to
verify plasmid integrity
thio plate
apr plate
Liquid culture
(no selection)
thio plate
apr plate
Liquid culture
(no selection)
thio plate
apr plate
Liquid culture
(no selection)
Replica plate
thio plate
apr plate
Master plate
(no selection)
Screen apr^R, thio^S clones
by hybridization

plasmid loss were monitored by plating aliquots of broth cultures on MYM agar plates containing apramycin or thiostrepton. When the sample plates indicated that the cultures were thiostrepton sensitive and apramycin resistant, aliquots were plated onto MYM agar without antibiotic. Isolated colonies were patched onto a master plate without selection and the master plates were used to replica plate MYM plates containing either thiostrepton or apramycin. Colonies that were apramycin resistant, thiostrepton sensitive were selected for further screening. Chromosomal DNA was isolated from several gene replacement isolates and probed with a ^{32}P-dCTP random primer-labelled apramycin resistance gene fragment (Figure III.5.3) and with similarly labelled *bldA* DNA. The apramycin resistance gene probe hybridized to a 2.8 kb *Sal*I fragment in all of the mutant strains whereas the *bldA* probe only hybridized to the 1.5 kb *Sal*I fragment in the wild-type strain. All isolates displayed a *bld* (for *bald)* phenotype but unexpectedly produced cephamycin C and clavulanic acid as determined by bioassay. There was some concern that the phenotype of the mutant strains could be due to a possible heterokaryon as the procedure used to isolate the mutants did not involve colony regeneration from a single cell or spore. Although the mutant strain is characterized as *bld* it can produce small quantities of spores at the periphery of well-isolated colonies on the appropriate medium (ISP#3) if incubated for several weeks. Spores were collected from two of the mutant strains and used to obtain new isolates from the regeneration of single spores. Again chromosomal DNA was isolated from the mutant strains and the replacement of the *bldA* gene was verified by Southern hybridization (Figure III.5.4). All of the mutants were confirmed to produce both cephamycin C and clavulanic acid by bioassay. Although the bioassays were not quantitative, it was apparent that antibiotic production in the mutant

Figure III.5.3

Hybridization analysis of *S. clavuligerus* Δ*bldA* mutants. Chromosomal DNA was isolated from Δ*bldA* mutants #4, #7, #19, #20, #25, #26, and #33. The DNA from the mutant strains and from *S. clavuligerus* NRRL 3585 (wild-type) strain was digested with *Sal*I, separated by gel electrophoresis, transferred to a nylon membrane, and probed with the gel-purified, ^{32}P-dCTP random primer-labelled *apramycin* gene (**A**) or ^{32}P-dCTP random primer-labelled *bldA* gene (**B**). The *apramycin* gene probe was purified as a 1.45 kb *Nco*I fragment, and the *bldA* probe was generated by PCR of p9S+ DNA using the primers NTR9 (5'-CCGCGAATTCGCCATGGAACGCCTTGT-3') which is homologous to a region upstream of the *bldA* gene, and NTR10 (5'-CGGAGCCGGACTCGAACC-3') which is complementary to the 3' end of the tDNA. Hybridization and washes were carried out at 65°C for the membrane probed with the *apramycin* gene and at 37 °C for the membrane probed with the *bldA* gene. The DNA size standard was λ DNA digested with *Pst*I.

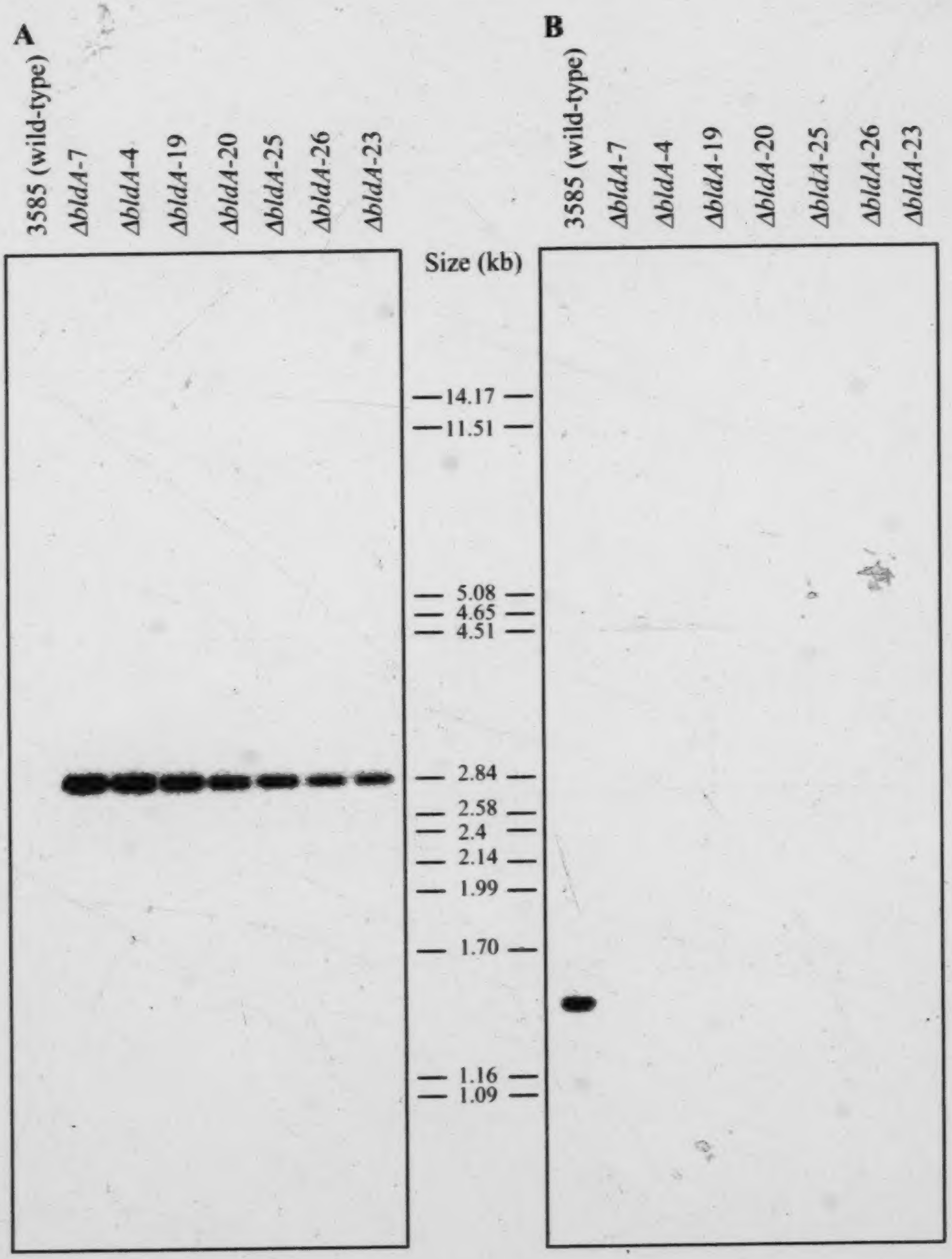
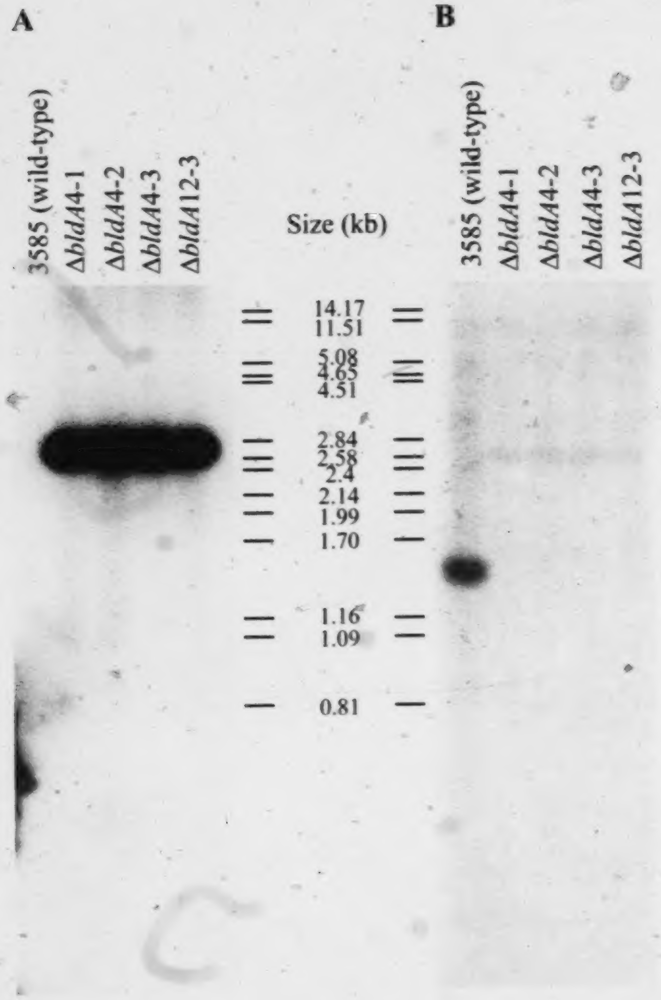

A
3585 (wild-type)
ΔbldA-7
ΔbldA-4
ΔbldA-19
ΔbldA-20
ΔbldA-25
ΔbldA-26
ΔbldA-23
Size (kb)
14.17
11.51
5.08
4.65
4.51
2.84
2.58
2.4
2.14
1.99
1.70
1.16
1.09
B
3585 (wild-type)
ΔbldA-7
ΔbldA-4
ΔbldA-19
ΔbldA-20
ΔbldA-25
ΔbldA-26
ΔbldA-23

Figure III.5.4

Hybridization analysis of *S. clavuligerus* Δ*bldA* mutants that were regenerated from spores. Chromosomal DNA was isolated from Δ*bldA* mutants #4-1, #4-2, #4-3, and #12-3. The DNA from the mutant strains and from *S. clavuligerus* NRRL 3585 (wild-type) was digested with *Sal*I, separated by gel electrophoresis, transferred to a nylon membrane, and probed with the gel-purified, ^{32}P-dCTP random primer-labelled *apramycin* gene **(A)** and the ^{32}P-dCTP random primer-labelled *bldA* gene **(B)**. The *apramycin* gene probe was purified from pUC120Ap(Nco) as a 1.45 kb *Nco*I fragment. The *bldA* probe was generated by PCR of p9S+ DNA using the primers NTR9 (5 '-CCGCGAATTCGCCATGGAACGCCTTGT-3') which is homologous to a region upstream of the *bldA* gene, and NTR10 (5'-CGGAGCCGGACTCGAACC-3') which is complementary to the 3' end of the tDNA. Hybridization and washes were carried out at 65°C for the *apramycin* probe and at 37°C for the *bldA* probe. The DNA size standard was λ DNA digested with *Pst*I.


A
B
3585 (wild-type)
Δ*bldA*4-1
Δ*bldA*4-2
Δ*bldA*4-3
Δ*bldA*12-3
Size (kb)
14.17
11.51
5.08
4.65
4.51
2.84
2.58
2.4
2.14
1.99
1.70
1.16
1.09
0.81

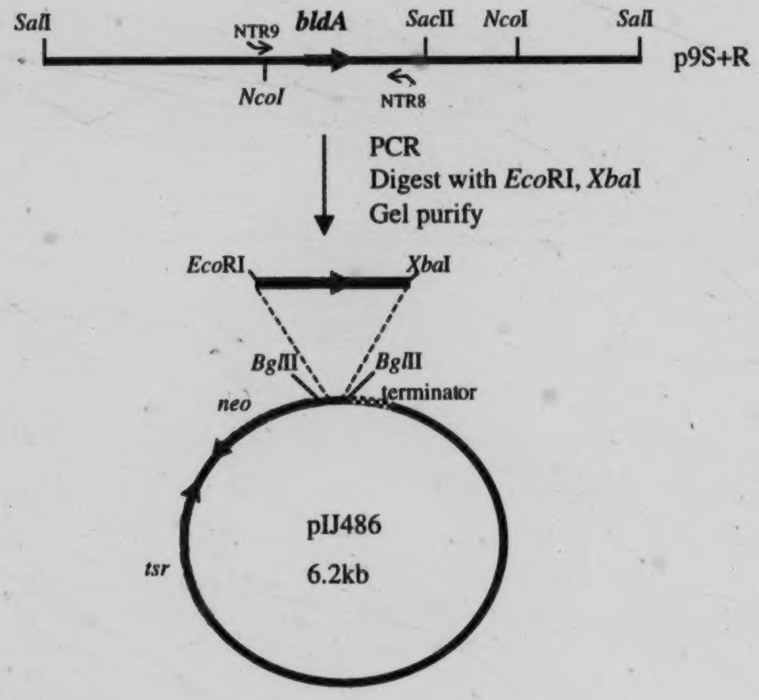
strains was not altered significantly as the zones of inhibition on the test plates using either supernatant or agar plugs from wild-type cultures or *bldA* mutant cultures were of similar size. Although all of the gene replacement isolates displayed an identical phenotype, a derivative of isolate 4 (isolate 4-1) was chosen for further study since isolate 4 (and it's derivatives) was never subjected to selection with apramycin whereas some of the other isolates were grown on the antibiotic.

III.6 COMPLEMENTATION OF THE *bldA* MUTANT STRAIN

In order to verify that the replacement of the *bldA* gene was responsible for the *bld* phenotype of the mutant strain, the wild-type gene was cloned back into the mutant strain. A 305 bp *bldA* fragment was generated by PCR using the primers NTR8 and NTR9 (Table II.1.1). NTR9 is homologous to the region 127-147 bases upstream from the tDNA where the *Nco*I restriction site is located. It also contains an *EcoR*I restriction site engineered into the 5' end. NTR8 is homologous to the region 45-64 bases downstream from the 3' end of the tRNA gene, just upstream of the *Sac*II restriction site, and contains an *Xba*I restriction site engineered into the 5' end. The amplified DNA fragment was expected to contain the promoter for the *bldA* gene based on homology to the *S. coelicolor bldA* promoter. The PCR product was cloned into *EcoR*I-*Xba*I digested pIJ486 (Figure III.6.1) and transformed into the *S. lividans bldA* strain J1725 (Leskiw *et al*., 1991b) prior to transforming into *S. clavuligerus*. The *S. lividans* J1725 strain is a *bldA* point mutant and was used as a host to modify the plasmid DNA prior to introducing it into *S. clavuligerus* as well as to determine if the *S. clavuligerus bldA* gene could complement an *S. lividans bldA* mutant. Introduction of the *S. clavuligerus bldA* gene on the high-copy number plasmid pIJ486 into the *S. lividans bldA* mutant strain

Figure III.6.1

Subcloning of the *S. clavuligerus bldA* gene into the *Streptomyces* vector pIJ486 for complementation of the *bldA* mutant phenotype. A 305 bp *bldA* fragment was generated by PCR using the primers NTR8 and NTR9. NTR9 (5'-CCGCGAATTCG**CCATGG**AACGCCTTGT-3') is homologous to the region 127-147 bases upstream from the tDNA where the *Nco*I restriction site (bold) is located and also contains an *Eco*RI restriction site (underlined) engineered into the 5' end. NTR8 (5'-CGCCTCTAGACGATGAGGCGATCTTGAA-3') is homologous to the region 45-64 bases downstream from the 3' end of the tRNA gene and contains an *Xba*I restriction site (underlined) engineered into the 5' end. The amplified DNA fragment was expected to contain the promoter for the *bldA* gene based on homology to the *S. coelicolor bldA* promoter. The PCR product was digested with *Eco*RI and *Xba*I and cloned into *Eco*RI-*Xba*I digested pIJ486.


SalI
NTR9
bldA
SacII
NcoI
SalI
p9S+R
NcoI
NTR8
PCR
Digest with EcoRI, XbaI
Gel purify
EcoRI
XbaI
BglII
BglII
terminator
neo
tsr
pIJ486
6.2kb

restored actinorhodin production as well as the ability to sporulate (Figure III.6.2). Similar results were observed when the plasmid was isolated from *S. lividans* J1725 and introduced into the *S. coelicolor bldA* deletion mutant strain J1681 (Leskiw *et al.*, 1991b) (data not shown). The *bldA* gene on the high-copy number plasmid also complemented the *S. clavuligerus bldA* mutant strain as it restored sporulation (Figure III.6.3, Panel A, B, and C). In order to verify that the *bldA* gene and not the plasmid vector restored sporulation, the plasmid pIJ486 without insert was also transformed into *S. clavuligerus* Δ*bldA* strain. As expected the vector without the insert containing the *bldA* gene did not complement the mutant phenotype (not shown). The presence of the *bldA* gene on the plasmid pIJ486 in the complemented *S. clavuligerus* Δ*bldA* transformants was verified by isolating plasmid DNA from complemented *bldA* mutants, digesting with *Bgl*II to release the cloned PCR fragment (see Figure III.6.1), separating on an agarose gel, transferring to a nylon membrane, and probing with the ^{32}P-labelled *bldA* gene (Figure III.6.4, lane pIJ486*bldA).* The *bldA* hybridizing fragment is larger than the 305 bp fragment inserted into pIJ486, approximately 550 bp, as the insert was released from the plasmid along with the entire multiple cloning site. Plasmid DNA was also isolated from *S. clavuligerus* Δ*bldA* transformants containing the vector without insert and subjected to the same procedure (Figure III.6.4, lane pIJ486). The membrane was stripped and probed with the ^{32}P-labelled *tsr* gene to verify that both transformants did contain the 6.2 kb pIJ486 plasmid (Figure III.6.5, lanes pIJ486*bldA* and pIJ486). In addition to the 6.2 kb bands observed there are fainter hybridizing bands approximately 4 kb in length in both pIJ486*bldA* and pIJ486, as well as an additional band approximately 5.8 kb in length in the pIJ486*bldA* lane alone. The most likely explanation for these additional hybridizing

Figure III.6.2

Complementation of the *S. lividans* J1725 *bldA* mutant phenotype by the *S. clavuligerus bldA* gene on the high copy number plasmid pIJ486. **(A)** *S. lividans* J1725 displaying the *bldA* phenotype on R2YE agar. **(B)** *S. lividans* J1725 (pIJ486*bldA*) on R2YE agar.

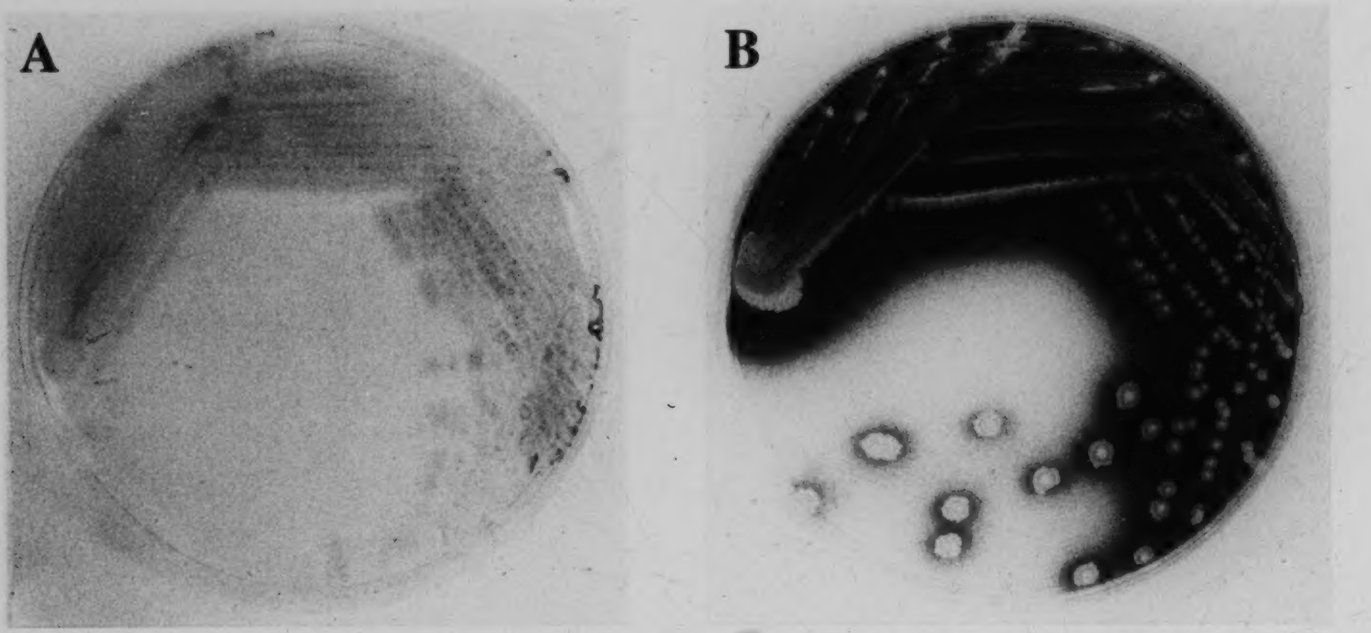

A
B

Figure III.6.3

Complementation of the *S. clavuligerus bldA* mutant phenotype.

(A) *S. clavuligerus* NRRL 3585 (wild-type) producing abundant grey-colored spores; **(B)** non-sporulating *S. clavuligerus ΔbldA* 4-1; **(C)** *S. clavuligerus ΔbldA* 4-1 (pIJ486*bldA*); **(D)** *S. clavuligerus ΔbldA* 4-1 (pSET152*bldA-tsr*). All the strains shown were grown on ISP#3 (sporulation) media.

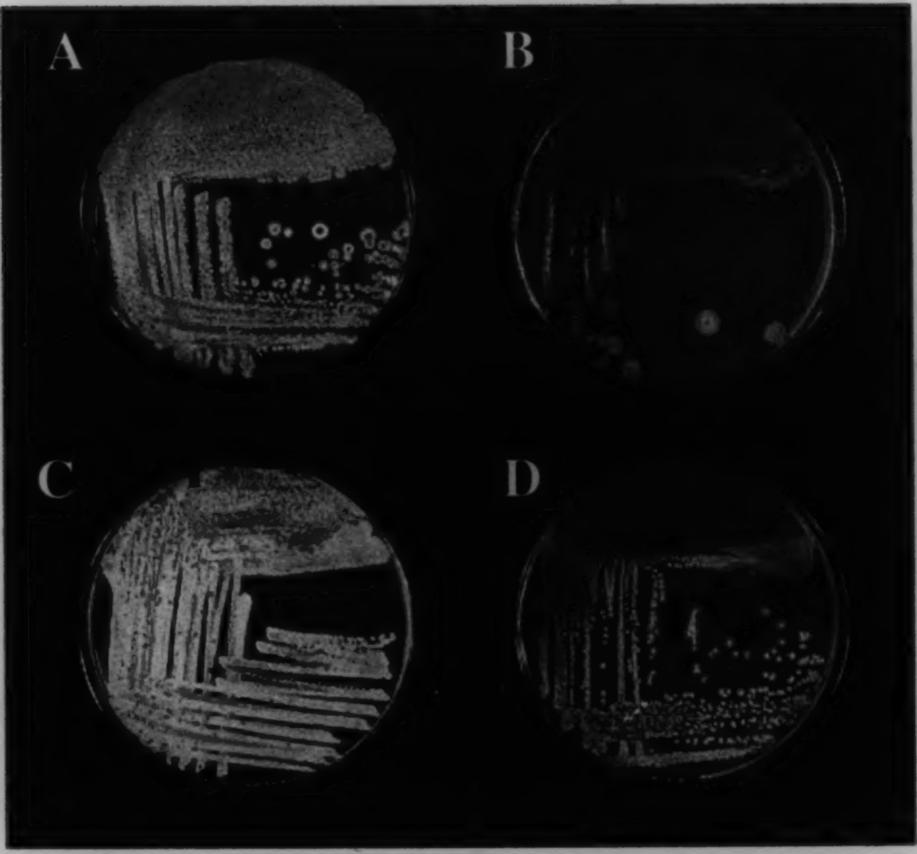

A
B
C
D

Figure III.6.4

Hybridization of DNA isolated from *S. clavuligerus* with the ^{32}P-labelled *bldA* gene. Chromosomal DNA was isolated from *S. clavuligerus* NRRL 3585 (wild-type), *S. clavuligerus* Δ*bldA*4-1, Δ*bldA*(pSET152*bldA-tsr*), Δ*bldA*(pSET152*tsr*). Plasmid DNA (pIJ486*bldA*) and (pIJ486) was isolated from *S. clavuligerus* Δ*bldA* strains containing the appropriate plasmids. The chromosomal DNA was digested with *Sal*I (S), *Eco*RI (E), or *Eco*RI-*Xba*I (E/X). The plasmid DNA was digested with *Bgl*II (B). The DNA was separated by gel electrophoresis, transferred to a nylon membrane and probed with the ^{32}P-dCTP random primer-labelled *S. clavuligerus bldA* gene. The *bldA* probe was generated by PCR of p9S+ DNA using the primers NTR9 (5'-CCGCGAATTCGCCATGGAACGCCTTGT-3') which is homologous to a region upstream of the *bldA* gene, and NTR10 (5'-CGGAGCCGGACTCGAACC-3') which is complementary to the 3' end of the tDNA. Hybridization and washes were carried out at 37°C. The molecular weight markers (*Pst*I-digested λ DNA) were detected by probing with random-primer, ^{32}P-labelled λ DNA. In this case, hybridization and washes were carried out at 65°C.

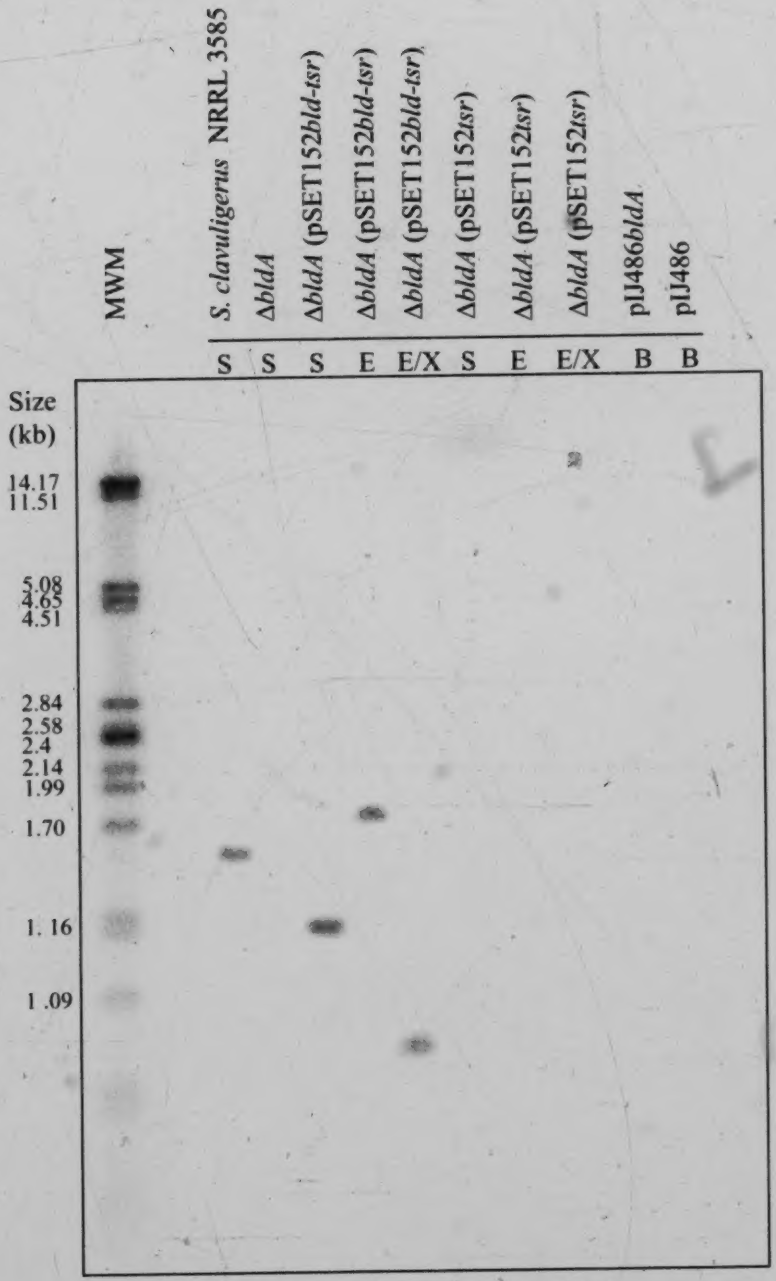
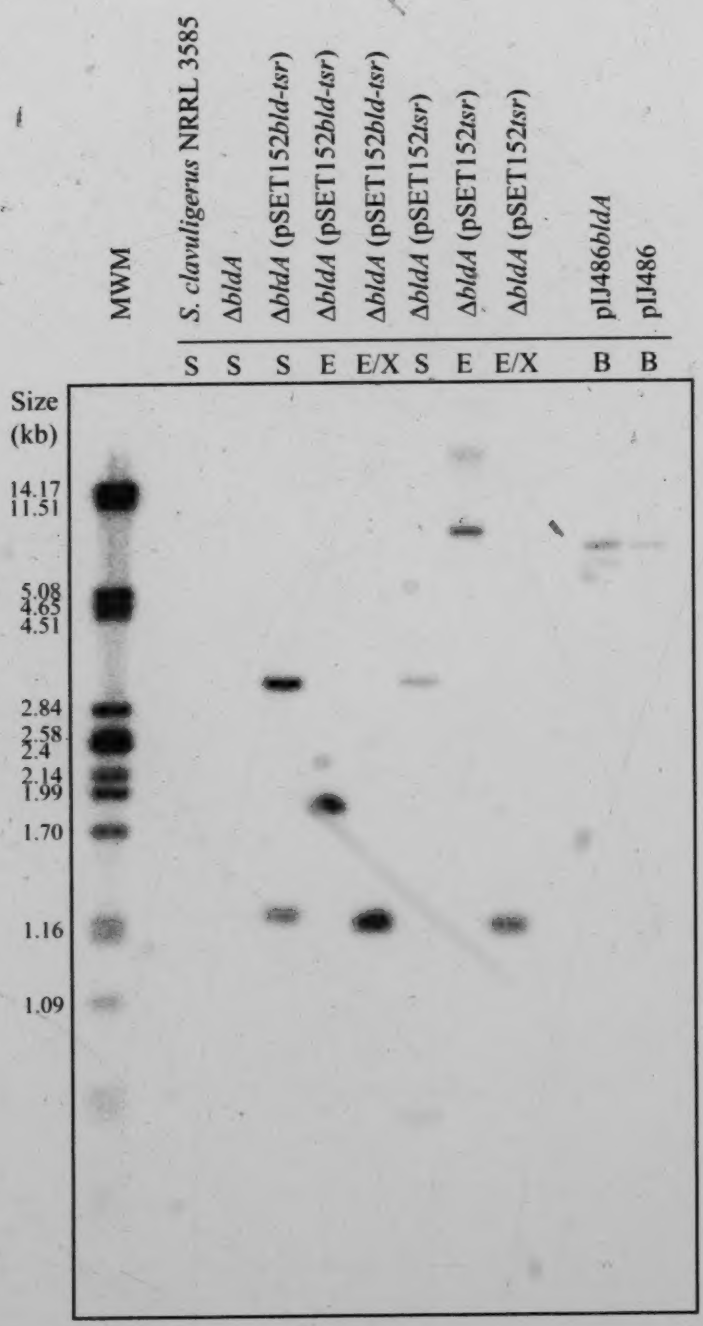

MWM
S. clavuligerus NRRL 3585
ΔbldA
ΔbldA (pSET152bld-tsr)
ΔbldA (pSET152bld-tsr)
ΔbldA (pSET152bld-tsr)
ΔbldA (pSET152tsr)
ΔbldA (pSET152tsr)
ΔbldA (pSET152tsr)
pIJ486bldA
pIJ486
S
S
S
E
E/X
S
E
E/X
B
B
Size (kb)
14.17
11.51
5.08
4.65
4.51
2.84
2.58
2.4
2.14
1.99
1.70
1. 16
1 .09

Figure III.6.5

Hybridization of DNA isolated from *S. clavuligerus* with the ^{32}P-labelled *tsr* gene. Chromosomal DNA was isolated from *S. clavuligerus* NRRL 3585 (wild-type), *S. clavuligerus* Δ*bldA*4-1, Δ*bldA*(pSET152*bldA-tsr*), Δ*bldA*(pSET152*tsr*). Plasmid DNA (pIJ486*bldA*) and (pIJ486) was isolated from *S. clavuligerus* Δ*bldA* strains containing the appropriate plasmids. The chromosomal DNA was digested with *Sal*I (S), *Eco*RI (E), or *Eco*RI-*Xba*I (E/X). The plasmid DNA was digested with *Bgl*II (B). The DNA was separated by gel electrophoresis, transferred to a nylon membrane and probed with the gel-purified, ^{32}P-dCTP random primer-labelled *tsr* gene. The *tsr* probe was isolated as an ~1 kb *EcoR*I-*Xba*I fragment from the plasmid pAU5 [Giebelhaus, 1996 #634]. Hybridization and washes were carried out at 65°C. The *Pst*I-digested λ DNA molecular weight markers were detected by probing with random-primer, ^{32}P-labelled λ DNA.

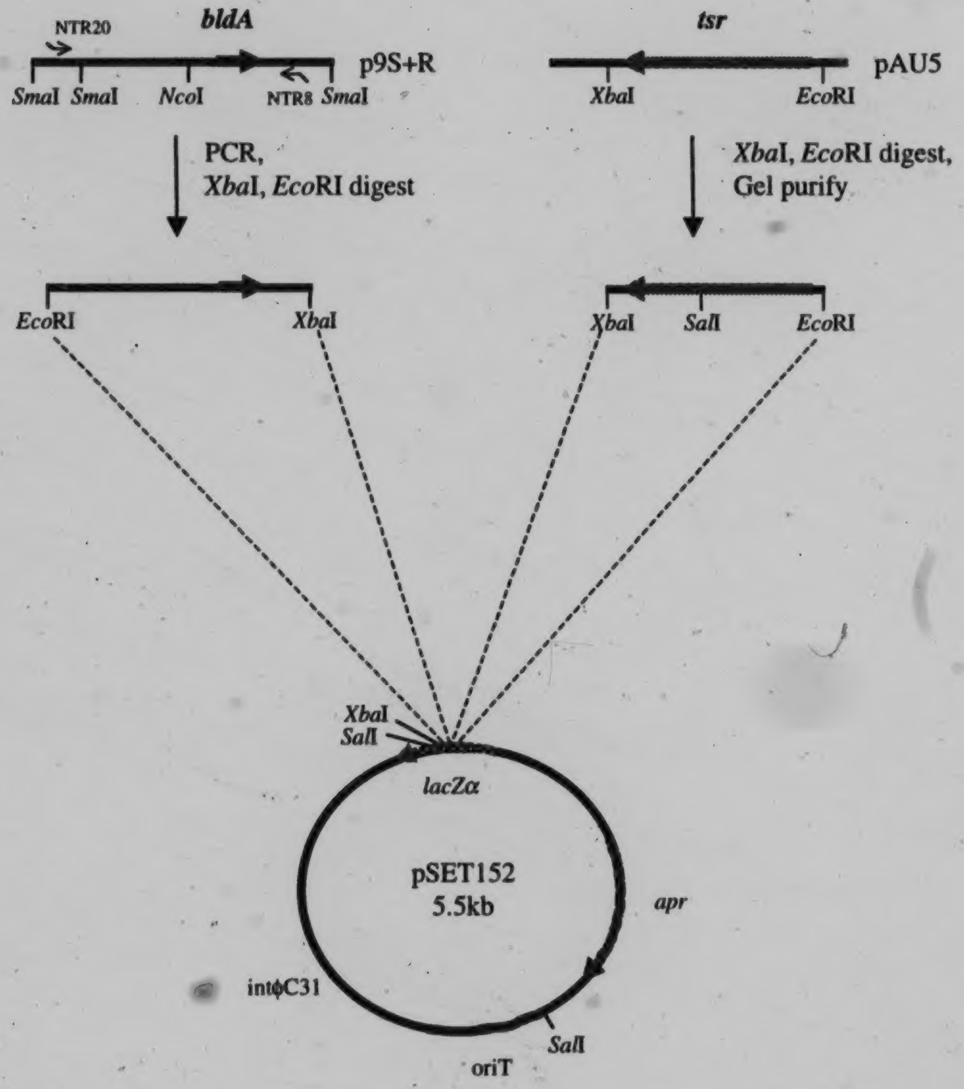

MWM
S. clavuligerus NRRL 3585
ΔbldA
ΔbldA (pSET152bld-tsr)
ΔbldA (pSET152bld-tsr)
ΔbldA (pSET152bld-tsr)
ΔbldA (pSET152tsr)
ΔbldA (pSET152tsr)
ΔbldA (pSET152tsr)
pIJ486bldA
pIJ486
S
S
S
E
E/X
S
E
E/X
B
B
Size (kb)
14.17
11.51
5.08
4.65
4.51
2.84
2.58
2.4
2.14
1.99
1.70
1.16
1.09

bands is that they are caused by undigested plasmid DNA (covalently closed circular) and partially digested plasmid DNA (relaxed circular). The results however, do verify that both transformed strains contain the plasmid pIJ486, as it hybridizes to the *tsr* probe, but only the complemented strain contains the *bldA* gene.

The *bldA* gene was also cloned into the *Streptomyces* integrating vector pSET152 to ensure that a single-copy of the *bldA* gene was sufficient to complement the mutant phenotype. This plasmid vector can be shuttled from *E. coli* to *Streptomyces* where it integrates into the ϕC31*att* site. A 665 bp fragment containing the *bldA* gene was amplified by PCR using the primers NTR20 and NTR8 (Table II.1.1). NTR8 hybridizes to the region 45-64 bases downstream from the 3' end of the tRNA gene, just upstream of the *Sac*II restriction site and contains an *Xba*I restriction site engineered into the 5' end. NTR20 hybridizes to the region approximately 490 nucleotides upstream of the tDNA and contains an *Eco*RI restriction site engineered at the 5' end. Additional sequence 5' of the *bldA* tDNA was included as it was determined that the *bldA* gene is transcribed by more than one promoter (discussed in section III.11-13). The NTR20 primer was believed at the time to be homologous to the 5' end of ORF1, which should be sufficiently upstream of the *bldA* tDNA to include all of the promoters. It was not until later that it was determined that the NTR20 primer is homologous to a region outside of the ORF1 sequence, so it is possible that the amplified *bldA* fragment doesn't contain all of the *bldA* promoter sequences. The amplified product was digested with *Eco*RI and *Xba*I for cloning into the *Streptomyces* integrating vector pSET152 (Bierman *et al.*, 1992). Since the selectable marker in pSET152 is the apramycin resistance gene and the *bldA* mutant strain was generated by replacement of the *bldA* gene with the apramycin resistance gene cassette it was necessary to clone an additional antibiotic resistance marker into the vector. The thiostrepton resistance marker was cloned into pSET152 as a 1.1 kb *Eco*RI-*Xba*I fragment after tail-to-tail ligation to the *Eco*RI-*Xba*I *bldA* gene fragment. The generalized scheme is shown in Figure III.6.6. The ligation mixture was transformed into *E. coli* DH5α using blue-white selection. Plasmid DNA was isolated from a number of transformants and digested with *Eco*RI to verify that the entire 1.8 kb tail-to-tail ligated *bldA* and *tsr* fragment was present, and with *Xba*I to determine the orientation of the *bldA-tsr* fragment. Plasmid isolates were obtained with the *bldA-tsr* fragment in both orientations, however, the orientation shown in Figure III.6.6 was chosen as the *bldA* gene is convergent to the *lacZ* gene so any expression of the *bldA* tRNA would be from its own promoter and not the result of read-through expression from the *lacZ* promoter. The pSET152(*bldA-tsr*) plasmid was transformed into *E. coli* ET12567, a *dam*$^-$, *dcm*$^-$ strain for generation of unmethylated plasmid DNA. The plasmid was then transformed into *S. clavuligerus* Δ*bldA* protoplasts and thiostrepton resistant transformants were selected. The pSET152 plasmid vector containing only the thiostrepton resistance marker pSET152(thio#2) was also transformed into *S. clavuligerus* Δ*bldA* as a negative control. The plasmid containing the *tsr* gene was generated by cloning the *tsr* marker as a *Bcl*I fragment into the unique *Bam*HI site of pSET152 (L. Geibelhaus, personal communication). The orientation of the *tsr* marker in pSET152(thio#2) was determined by restriction analysis and was found to be in the same orientation as the *tsr* gene in pSET152*(bldA-tsr)*. Integration of the plasmids pSET152*(bldA-tsr)* and pSET152(thio#2) into the ϕC31*att* site of *S. clavuligerus* Δ*bldA*(4-1) was verified by probing chromosomal DNA with the *bldA* gene

Figure III.6.6

Subcloning of the *S. clavuligerus bldA* gene into the integrating *Streptomyces* vector pSET152 for complementation of the *bldA* mutant phenotype. A 665 bp fragment containing the *bldA* gene was amplified by PCR using the primers NTR20 (5'-GCTAGAATTCGGAGGCCGTCTCCGAGGA-3') and NTR8 (5'-CGCCTCTAGACGATGAGGCGATCTTGAA-3'). NTR20 is homologous to the region approximately 490 nucleotides upstream of the *bldA* tDNA and contains an *Eco*R1 restriction site (underlined) engineered at the 5' end. NTR8 is homologous to the region 45-64 bases downstream from the 3' end of the tRNA gene and contains an *Xba*I restriction site (underlined) engineered into the 5' end. The amplified product was digested with *Eco*RI and *Xba*I, ligated tail-to-tail with a similarly digested fragment containing the *tsr* gene, and cloned into *Eco*RI-digested pSET152. The *tsr* gene was gel-purified from the plasmid pAU5 (Giebelhaus *et al.*, 1996) after digestion with *Eco*RI-*Xba*I. The orientation of the *bldA-tsr* insert was determined by restriction analysis.


NTR20
bldA
p9S+R
SmaI
SmaI
NcoI
NTR8
SmaI
tsr
pAU5
XbaI
EcoRI
PCR,
XbaI, EcoRI digest
XbaI, EcoRI digest,
Gel purify
EcoRI
XbaI
XbaI
SalI
EcoRI
XbaI
SalI
lacZα
pSET152
5.5kb
apr
intφC31
SalI
oriT

(Figure III.6.4) and the *tsr* gene (Figure III.6.5). As expected, the *bldA* probe only hybridized to chromosomal DNA from *S. clavuligerus* NRRL 3585 (wild-type) and *S. clavuligerus ΔbldA* (pSET152*bldA-tsr*), and the *tsr* probe only hybridized to *S. clavuligerus* strains containing the pSET152*bldA-tsr* or pSET152*tsr* vector. The probes hybridized to fragments of the expected size (Table III.6.1), although the fragments in the 650-700 bp range appear slightly larger, which is likely the result of high salt concentrations in the restriction digestions retarding the migration of these small fragments. As the *ΔbldA*(pSET152*tsr)* vector only contains a single *Eco*RI site, the large ~9 kb hybridizing fragment observed is expected to contain the *tsr* gene, pSET152 sequence and chromosomal DNA *attB* sequence. The results of the southern analysis verify that the *bldA* gene in the plasmid vector pSET152 is integrated into the *attB* site on the *S. clavuligerus* chromosome, and is the only copy of the *bldA* gene in the mutant strain. Complementation analysis indicated that the single copy of the *bldA* gene was sufficient to complement the *bldA* phenotype (Figure III.6.3, Panel D), whereas insertion of the vector alone did not restore sporulation (data not shown). The *S. clavuligerus ΔbldA*(pSET152*bldA-tsr)* strain did not sporulate as vigorously as the wild-type strain or the strain complemented with the *bldA* gene on a high copy vector. This could be the result of a positional effect as the single-copy *bldA* gene is integrated at the ϕC31*att* site and not at it's native location, or could indicate that the 665 bp *bldA* fragment doesn't contain all of the *bldA* promoters.

III.7 IDENTIFICATION OF THE *S. coelicolor leuU* tRNA GENE

Previous analysis of the expression of the *S. coelicolor bldA* tRNA compared the temporal accumulation of the mature *bldA* tRNA transcript with the accumulation of the

Table III.6.1 Expected size of hybridizing *S. clavuligerus* chromosomal DNA fragments to the *bldA* and *tsr* probe

Size of *S. clavuligerus* chromosomal DNA restriction fragments expected to hybridize to the *bldA* probe

Restriction Enzyme	3585 (wild-type)	Δ*bldA*	Δ*bldA* (pSET152*bld-tsr)*	Δ*bldA* (pSET152*tsr)*
Sal I	1.5 kb	-	1.3 kb	-
Eco RI	N/D	-	1.8 kb	-
Eco RI/*Xba* I	N/D	-	700 bp	-

Size of *S. clavuligerus* chromosomal DNA restriction fragments expected to hybridize to the *tsr* probe

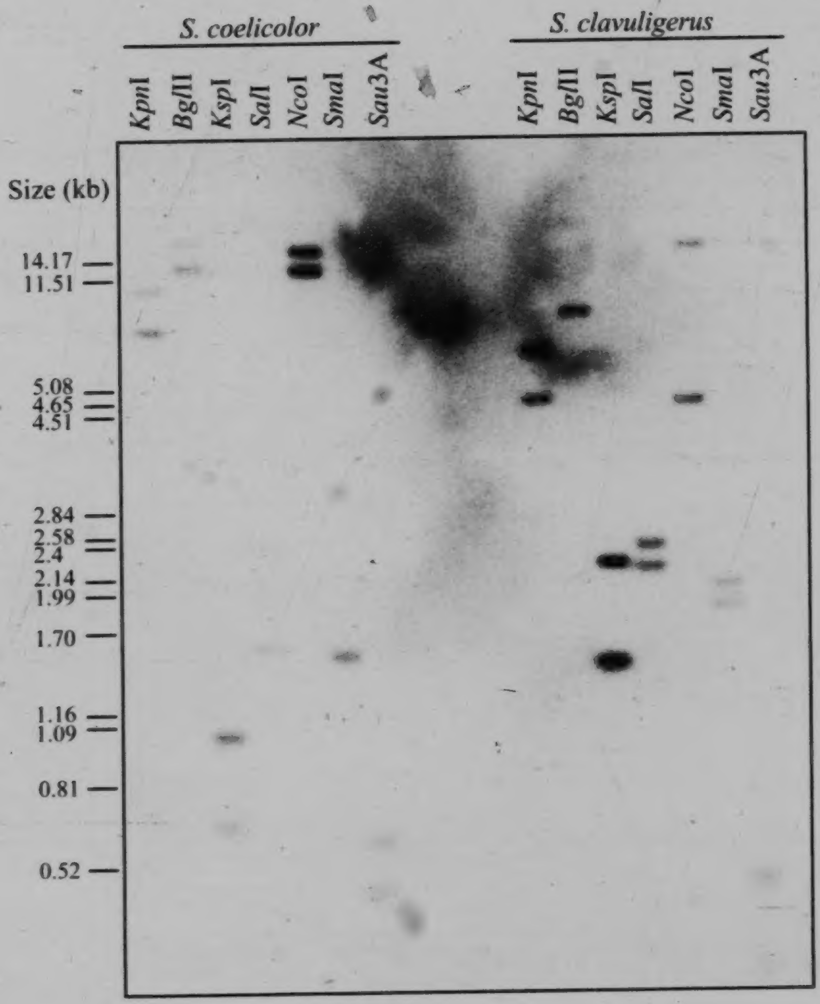
Restriction Enzyme	3585 (wild-type)	Δ*bldA*	Δ*bldA* (pSET152*bld-tsr)*	Δ*bldA* (pSET152*tsr)*
Sal I	-	-	3 kb	3 kb
	-	-	1.3 kb	650 kb
Eco RI	-	-	1.8 kb	N/D
Eco RI/*Xba* I	-	-	1.2 kb	1.2 kb

N/D Not determined or unknown

lysTα-specified tRNA (Leskiw *et al.*, 1993). In constrast to the temporal pattern of accumulation demonstrated by the *bldA* tRNA, the major lysyl-tRNA ($\mathrm{tRNA}^{\mathrm{Lys}}_{\mathrm{CTT}}$) was expressed at relatively constant levels throughout growth. While this relatively constant level of expression is expected for the majority of tRNAs involved in protein synthesis, a more suitable control would be to compare the expression of the *bldA*-leucyl tRNA with the expression of a major leucyl-tRNA. Since the *bldA* tRNA genes were the only leucyl-tRNA genes identified in *Streptomyces* at the time, and since *Streptomyces* tRNA genes show significant similarity to *E. coli* tRNAs, leucyl-tRNA sequences from *E. coli* were compared and used to design degenerate oligonucleotide primers for the amplification of *Streptomyces* leucyl-tRNA genes. The primers used were 16-mer, 8-fold degenerate, synthetic oligonucleotides NTR6 (5'-$\mathrm{GC}^{\mathrm{C}}_{\mathrm{G}}$ $\mathrm{G}^{\mathrm{GG}}_{\mathrm{AA}}$ GTGGCGGAAT-3'), homologous to the 5' end of *E. coli* leucyl-tRNAs, and NTR7 (5'-$^{\mathrm{GG}}_{\mathrm{AA}}$ $\mathrm{GGC}^{\mathrm{C}}_{\mathrm{G}}$ $\mathrm{GGACT}^{\mathrm{T}}_{\mathrm{C}}$ GAAC-3'), complementary to the 3' end. PCR amplification was carried out on *S. coelicolor* J1681 (Δ*bldA*;(Leskiw *et al.*, 1993)) genomic DNA and the major amplification product was blunt-end cloned into the phage vector M13mp19 and sequenced. Using this methodology a potential *S. coelicolor* leucyl-tRNA containing a GAG anticodon was identified. Since the sequence at the ends of the amplified DNA corresponded to the degenerate primers rather than to the *S. coelicolor* tRNA gene sequence, the internal tRNA sequence was used to design a second set of primers (BKL42 and BKL43) for inverse PCR amplification (Ochman *et al.*, 1990) and Southern hybridization. Southern hybridization was carried out on chromosomal DNA digests of *S. coelicolor* J1681 with the ^{32}P-labelled BKL42 (5'-CTCAAGCTAGCGCGTCTG-3') primer (which is complementary to the internal tRNA gene sequence from the anticodon to the D-loop) to identify suitable sized fragments for inverse PCR amplification. Hybridization and washing were carried out at 50°C under conditions allowing only a single mismatch. Under these condition the primer is predicted to hybridize to only CUC translating leucyl-tDNAs as other leucyl-tDNA species would be expected to have a minimum of three mismatches. This prediction is based on the observed homology of *E. coli* leucyl-tRNAs over this region. One or two hybridizing bands were detected in each digest (Figure III.7.1) suggesting that there are at least two copies of CUC translating leucyl-tRNA genes in the *S. coelicolor* genome. *S. clavuligerus* chromosomal DNA also showed one or two hybridizing bands when probed with BKL42 under the same conditions indicating that the leucyl-tRNA genes in these two species are highly conserved and that there are also at least two copies of a CUC translating leucyl-tRNA gene in the *S. clavuligerus* genome. *S. coelicolor* chromosomal DNA digested with *Ksp*I (*Sst*II), which showed hybridizing bands of 1 kb and 670 bp, and *S. clavuligerus* chromosomal DNA digested with *Ksp*I which showed hybridizing bands of 1.6 kb and 2.3 kb, was ligated under conditions that favor intramolecular ligation and used as template for inverse PCR. Inverse PCR was carried out with the primers BKL42 (5'-CTCAAGCTAGCGCGTCTG-3') and BKL43 (5'-GTGCTAGTGCCCTTTATC-3'). BKL43 is homologous to the putative CUC translating leucyl-tDNA from the anticodon loop to the variable loop. The 670 bp and 1 kb *S. coelicolor*, and 1.6 kb *S. clavuligerus* inverse PCR products were purified and sequenced directly by a sequencing service. The flanking sequence obtained from each of the inverse PCR products was used to design additional primers to amplify the intact tDNA. The primers designed from the 1 kb *S. coelicolor Ksp*I sequence were: BKL48 (5'-CACTGCTGGTGGCTCCC-3'), which is homologous to sequence upstream from the

Figure III.7.1

Hybridization of chromosomal DNA from *S. coelicolor* J1501 and *S. clavuligerus* NRRL 3585 with the ^{32}P-labelled primer BKL42 to detect leucyl-tRNAs with the anticodon GAG. The chromosomal DNA was digested with *Kpn*I, *Bgl*II, *Ksp*I, *Sal*I, *Nco*I, *Sma*I, and *Sau*3A. The DNA fragments were separated by gel electrophoresis, transferred to a nylon membrane and probed with the end-labelled BKL42 (5'-CTCAAGCTAGCGCGTCTG-3') primer. Hybridization and washes were carried out at 50°C. The DNA size standard was λ DNA digested with *Pst*I.


S. coelicolor
S. clavuligerus
KpnI
BglII
KspI
SalI
NcoI
SmaI
Sau3A
KpnI
BglII
KspI
SalI
NcoI
SmaI
Sau3A
Size (kb)
14.17
11.51
5.08
4.65
4.51
2.84
2.58
2.4
2.14
1.99
1.70
1.16
1.09
0.81
0.52

leucyl-tRNA gene; and BKL49 (5'-CTTCTCGTGCGACTCGG-3') which is complementary to sequence downstream from the leucyl-tRNA gene. These primers were used to amplify the intact leucyl-tRNA and some flanking sequence from *S. coelicolor* genomic DNA. The PCR amplification product was blunt-end cloned into pUC119 and sequenced by a sequencing service (Figure III.7.2). The primers designed from the 670 bp *S. coelicolor* *Ksp*I inverse PCR product were NTR11 (5'-GCCGAATTCACCTGGACCACCCACAAG-3') and NTR12 (5'-GCCAAGCTTCACCGTATTTTCGCCCGT-3'). NTR11 is homologous to sequence upstream of the leucyl-tRNA gene and contains an *Eco*RI restriction site (underlined) at the 5' end, NTR12 is complementary to sequence downstream of the leucyl-tRNA gene and contains a *Hin*dIII restriction site (underlined) at the 5' end. The restriction sites were engineered into the 5' end of these primers to allow for directional cloning of the PCR product. However, PCR with these primers failed to yield the expected product of 325 bp. Additional PCR using various combinations of primers NTR11, NTR12, BKL42 and BKL43 indicated that the NTR11 primer doesn't hybridize to the expected region. The reason for the failure of the NTR11 primer to amplify the leucyl-tRNA later became evident when the complete sequence of this tRNA gene was obtained during the course of sequencing cosmid 8A6 (Genbank Accession # AL031013) by the *Streptomyces coelicolor* sequencing project (Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge). One of the *Ksp*I restriction sites, which was used to circularize the 670 bp leucyl-tRNA containing fragment prior to performing inverse PCR is located 32 bp upstream from the 5' end of the tDNA. The primer NTR11 was designed from inverse PCR sequence past this *Ksp*I restriction site and corresponds to sequence 467-487 bases downstream from the 3' end of the tDNA. The *Ksp*I restriction site was not evident

Figure III.7.2

Nucleotide sequence of the intact, 325 bp, leucyl tRNA-encoding DNA fragment. The primers used to amplify the fragment are indicated by solid lines with half arrowheads. NTR6 (5'-GC$^{C}_{G}$ G$^{GG}_{AA}$ GTGGCGGAAT-3') and NTR7 (5'-$^{GG}_{AA}$ GGC$^{C}_{G}$ GGACT$^{T}_{C}$ GAAC-3') are the degenerate primers used in the initial PCR amplification and mismatches are indicated by ^. BKL42 (5'-CTCAAGCTAGCGCGTCTG-3') and BKL43 (5'-GTGCTAGTGCCCTTTATC-3') are the primers used for inverse PCR amplification of the tDNA and flanking sequence on the 1 kb *S. coelicolor Ksp*I fragment (BKL42 was also used for Southern hybridization). BKL48 (5'-CACTGCTGGTGGCTCCC-3') and BKL49 (5'-CTTCTCGTGCGACTCGG-3') are the primers used to amplify the tDNA and its flanking sequences. The mature form of the tRNA$^{Leu}_{GAG}$ is indicated in grey text and the anticodon is marked with (***). An inverted repeat which is located immediately downstream from the tRNA coding sequence is indicated by unlabelled, dashed arrows.

BKL48
CACTGCTGGTGGCTCCCCCTCATCCTCCCCCCGTGGCCGTCCGGACCCAA 50
GTGACGACCACCGAGGGGGAGTAGGAGGGGGGCACCGGCAGGCCTGGGTT

GAAGGGGGTGATCGGCCACCCCGGAAAGCCTGGTAATGTTTACGTCGTCG 100
CTTCCCCCACTAGCCGGTGGGGCCTTTCGGACCATTACAAATGCAGCAGC

NTR6
CCAAGGGGAACACCCCACGCGACAGACACCTTGTCCGGGTGGCGGAATGG 150
GGTTCCCCTTGTGGGGTGCGCTGTCTGTGGAACAGGCCCACCGCCTTACC

*** BKL43
CAGACGCGCTAGCTTGAGGTGCTAGTGCCCTTTATCGGGCGTGGGGGTTC 200
GTCTGCGCGATCGAACTCCACGATCACGGGAAATAGCCCGCACCCCCAAG
BKL42

AAGTCCCCCCTCGGACACCAGCTGAAACCCCTGCTGAGCAGGGGTTTTCT 250
TTCAGGGGGGAGCCTGTGGTCGACTTTGGGGACGACTCGTCCCCAAAAGA
NTR7

GCTTTTCCTGCACCGTGTCTCCCCTACAGTGGTCACCATGCCCACCCCTC 300
CGAAAAGGACGTGGCACAGAGGGGATGTCACCAGTGGTACGGGTGGGGAG

CCTGCCCCTGCGGGCGGTCCGAGTCGCACGAGAAG 335
GGACGGGGACGCCCGCCAGGCTCAGCGTGCTCTTC
BKL49

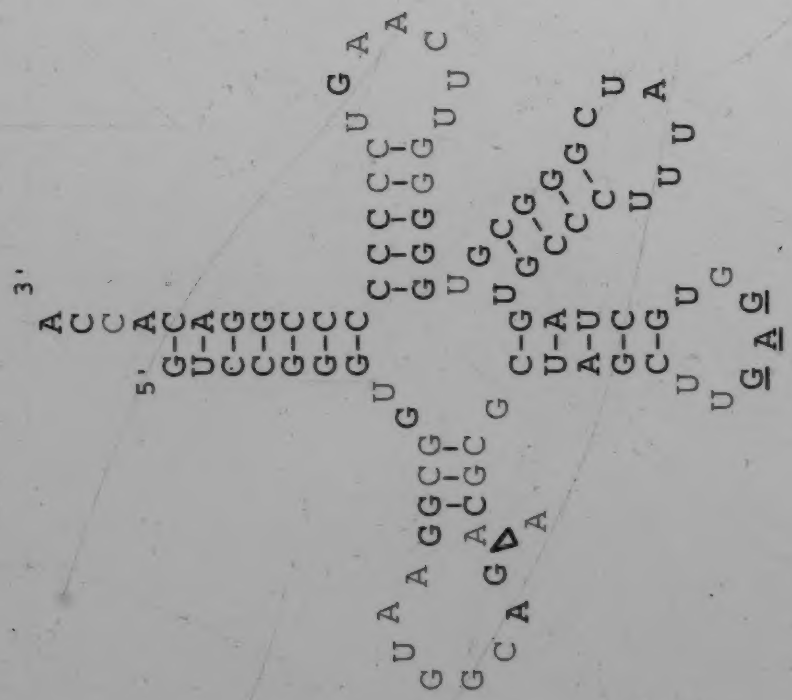
in the inverse PCR sequence as the sequence in this region contained several errors.

The sequence obtained from the 1.6 kb *S. clavuligerus Ksp*I fragment downstream of the tRNA was of poor quality and could not be used to design an additional primer. As a result, the only intact tRNA gene sequence obtained by this method was generated from the 1 kb *S. coelicolor Ksp*I fragment.

The DNA sequence of the 325 bp cloned fragment containing the leucyl-tRNA gene (Figure III.7.2) revealed that the tDNA is 88 nucleotides, and that it shares significant homology to *E. coli* leucyl-tRNAs. The sequence shows 75% identity to *E. coli* $\text{tRNA}^{\text{Leu}}_{\text{GAG}}$ and 83% identity to *E. coli* $\text{tRNA}^{\text{Leu}}_{\text{CAG}}$, while it shares only 65% identity with *S. coelicolor bldA* tRNA. The deduced secondary structure of the *S. coelicolor* $\text{tRNA}^{\text{Leu}}_{\text{GAG}}$ is shown in Figure III.7.3. It has been named *leuU* (Trepanier *et al.*, 1997) based on conventional nomenclature for prokaryotic tRNAs (Fournier and Ozeki, 1985). Since a second *leuU* tRNA gene in *S. coelicolor* has now been identified (Genbank Accession #AL031013, Table III.4.3), the first *leuU* gene becomes *leuUα*, and the second *leuUβ*. The *leuUα* tRNA contains a variable loop of 17 bases, a feature typical of Class II tRNAs which encode large variable loops. Unlike the majority of *Streptomyces* tRNA genes, including *leuUβ*, the 3' CCA terminus is encoded (see Table III.4.1). The D-loop of both *leuU* tRNAs contains the highly conserved G18 nucleotide, which is present in all *Streptomyces* tRNAs except the *bldA* tRNAs. The invariant A21 is absent from both *leuU* tRNAs. This base is also absent from the *bldA* leucyl-tRNAs indicating that the absence of this nucleotide may be a common feature of leucyl-tRNAs in *Streptomyces*. The *leuUα* and *leuUβ* tRNAs are virtually identical, the only difference between the two genes is the C75 A76 nucleotides which are not encoded in *leuUβ*. The *leuU* tRNAs

Figure III.7.3

Deduced secondary structure of the *S. coelicolor leuUα* tRNA. The anticodon sequence is underlined and the *Streptomyces* invariant and semi-invariant bases are shown in grey. The missing invariant A21 is also indicated in grey.



S. coelicolor leuUα

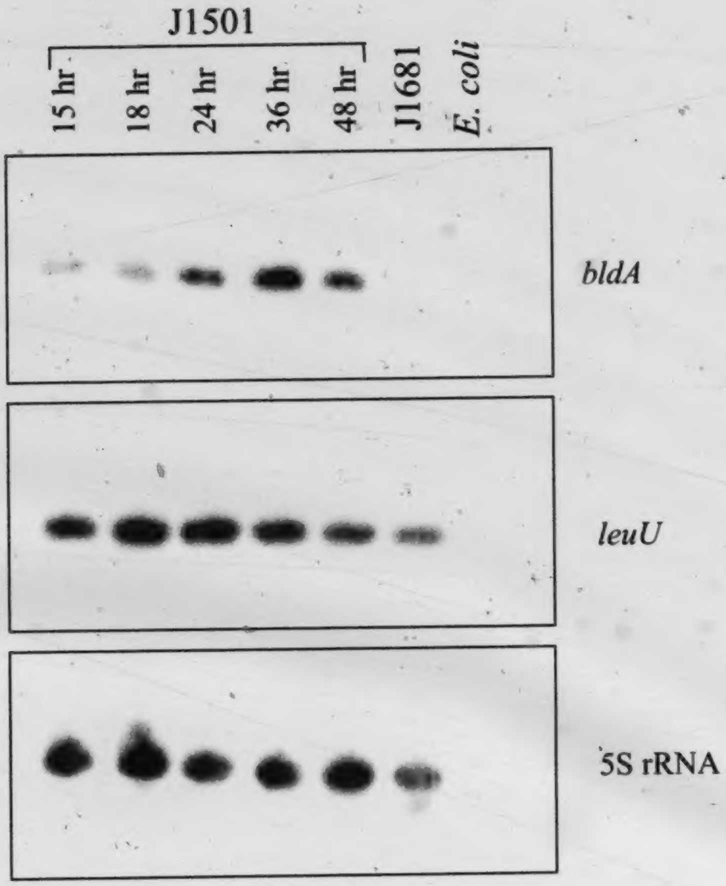
encode a GAG anticodon which translates CUC codons in mRNA, and according to codon usage, this is the second most abundant leucine codon in *Streptomyces* genes, accounting for 38.5% of leucine codons (Wright and Bibb, 1992). As this is a major leucyl-tRNA, it should be relatively abundant throughout growth, particularly during vegetative growth when the levels of protein synthesis are expected to be high. Since the two *leuU* sequences are almost identical, Northern analysis of $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{GAG}}$ with the primer BKL42 will detect transcripts originating from both tRNA genes. Analysis of the *leuUα* flanking sequence revealed an inverted repeat ($\Delta G = -19.2$ kcal) immediately downstream of the tRNA followed by a string of T residues which could be a rho-independent terminator. There is also a putative promoter sequence (based on comparison of *Streptomyces* promoter sequences (Bourn and Babb, 1995)) 43-51 bases upstream from the tDNA.

III.8 NORTHERN ANALYSIS OF THE *S. coelicolor bldA* AND *leuU* TRANSCRIPTS

Since the mature, processed form of the *S. coelicolor bldA* tRNA shows an unusual pattern of accumulation, with significant amounts of the tRNA being present only late in growth (Leskiw *et al.*, 1993), Northern blot analysis of the *bldA* and *leuU* tRNAs was compared to confirm that this late accumulation is a feature unique to *bldA* and not a general feature of leucyl-tRNAs in *Streptomyces* (Figure III.8.1). RNA was isolated from surface-grown cultures of *S. coelicolor* J1501 at 15, 18, 24, 36, and 48 hours post inoculation and subjected to Northern blot analysis. RNA isolated from 15 and 18 hour samples are from vegetative (preantibiotic) cultures. At 24 hours the cultures had begun to produce antibiotics and aerial hyphae. RNA isolated from *S. coelicolor* J1681 (Δ*bldA*) was used as a negative control for *bldA* transcripts, and

Figure III.8.1

Northern blot analysis of the *S. coelicolor bldA* tRNA and *leuU* tRNA in RNA samples (5 μg) isolated at various times (hours post inoculation) from surface-grown cultures of *S. coelicolor* J1501. The 15 and 18 hour samples are from vegetative (preantibiotic) cultures. At 24 hours the cultures had begun to produce antibiotics and aerial hyphae. Controls included *E. coli* RNA and RNA isolated at 36 hours from the *bldA* deletion strain, *S. coelicolor* J1681. *bldA* transcripts were detected by hybridization at 50°C with the ^{32}P-labelled oligonucleotide BKL5 (5'-TTAAGCTCGCCGTGTCT-3'); *leuU* transcripts were detected by hybridization at 56°C with the ^{32}P-labelled oligonucleotide BKL42 (5'-CTCAAGCTAGCGCGTCTG-3'); and the 5S rRNA transcripts were detected by hybridization at 60°C with the ^{32}P-labelled oligonucleotide BKL53 (5'-CCCTGCAGTACCATCGGCGCT-3').


J1501
15 hr
18 hr
24 hr
36 hr
48 hr
J1681
E. coli
bldA
leuU
5S rRNA

E. coli RNA was used as a negative control for *leuU* and 5S rRNA transcripts. Duplicate sets of RNA (5 µg/sample) were denatured for 10 minutes at 90°C in formamide dye mix (Sambrook *et al.*, 1989) and then separated on an 8% denaturing polyacrylamide gel. pBR322 DNA (250 µg), digested with *Hae*III, was treated in a similar manner and used as size standard. The samples were transferred to a nylon membrane and probed at 50°C with a ^{32}P-end-labelled oligonucleotide, BKL5 (5'-TTAAGCTCGCCGTGTCT-3'), which is complementary to the *bldA* tRNA from the anticodon to the D-loop, to detect *bldA* transcripts. The membrane was then stripped and probed at 56°C with a ^{32}P-end-labelled oligonucleotide, BKL42 (5'-CTCAAGCTAGCGCGTCTG-3'), which is complementary to the *leuUα* and *leuUβ* tRNAs from the anticodon to the D-loop, to detect *leuU* transcripts. Since these hybridizations were performed at 2°C below the T_d for each of the oligonucleotides, only homologous sequences will have hybridized to the probes. *E. coli* $\text{tRNA}^{\text{Leu}}_{\text{UAA}}$ is not detected with the BKL5 probe as seven mismatches would have to be allowed. The *E. coli* $\text{tRNA}^{\text{Leu}}_{\text{GAG}}$ and other leucyl-tRNAs are not detected with the BKL42 probe as a minimum of three mismatches would have to be allowed. As a control for RNA loading, the membrane was also probed with a ^{32}P-end-labelled 21-mer, synthetic oligonucleotide BKL53 (5'-CCCTGCAGTACCATCGGCGCT-3') at 60°C to detect *S. coelicolor* 5S rRNA transcripts. The *E. coli* 5S rRNA transcripts are not detected under these conditions as three mismatches would have to be allowed for the 5S rRNA transcripts to bind to the probe BKL53. The presence of *E. coli* RNA in approximately equivalent amounts was verified by staining and photographing the gel prior to transferring the RNA to a nylon membrane. Consistent with previous reports, the mature *bldA* tRNA showed a temporal pattern of accumulation, it is detectable at low levels during vegetative growth (15-18 hr) and increases during differentiation. In contrast, the transcript corresponding to the expected size of the processed *leuU* tRNA was present at relatively high levels during all stages of growth. The presence of the *leuU* tRNA throughout growth is typical of tRNAs in general, where they are expected to be expressed at abundant levels especially during early growth when levels of protein synthesis are high.

III.9 ANALYSIS OF ANTIBIOTIC PRODUCTION AND *bldA* tRNA EXPRESSION IN *S. clavuligerus*

Since it was demonstrated that the *bldA* tRNA shows a temporal pattern of accumulation in *S. coelicolor* (Leskiw *et al.*, 1993) it was of interest to determine if the *bldA* tRNA would show the same pattern of accumulation in *S. clavuligerus*. RNA was isolated (in duplicate) from surface-grown cultures of *S. clavuligerus* NRRL 3585 (wild-type) and *bldA* mutant strains at 30, 36, 48, 60, 72, 96, 120, 144, and 160 hours post-inoculation. Antibiotic production was determined by bioassay at each of the time-points (Table III.9.1). The wild-type strain (*S. clavuligerus* NRRL 3585) was just starting to produce cephamycin C at 36 hours, as one sample produced a slight zone of inhibition and the duplicate sample did not, however clavulanic acid was not detected until 48 hours. The zones of inhibition on the clavulanic acid test plates from 72-168 hours are in part due to the presence of one or more of the antibiotics holomycin, tunicamycin and cephamycin C, as the control plates for these time-points indicate that these two antibiotics are dected from 72-96 hours on. The *bldA* strain did not produce either cephamycin C or clavulanic acid until the 48 hour time-point at which time both cephamycin C and clavulanic acid were detected. Holomycin, tunicamycin and cephamycin C were detected in the control plates at 96 hours, so the increase in the zone

Table III.9.1 Antibiotic Bioassay of Duplicate Surface-grown Cultures of *S. clavuligerus*

S. clavuligerus NRRL 3585 (wild-type)

Sample Time	Zone of inhibition (diameter in mm)								
	Cephamycin C			Clavulanic Acid[2]					
	#1	#2	Average	test #1	test #2	test average	control #1	control #2	control average
30	-	-	-	-	-	-	-	-	-
36	slight[1]	-	slight	-	-	-	-	-	-
48	17	15	16	15	13	14	-	-	-
60	17	18	18	17	17	17	-	-	-
72	21	19	20	19	19	19	slight	-	4
96	26	25	26	28	21	25	12	slight	10
120	25	25	25	27	25	26	15	15	15
144	22	24	23	27	25	26	14	15	15
168	23	23	23	27	25	26	15	17	16

S. clavuligerus Δ*bldA* 4-1

Sample Time	Zone of inhibition (diameter in mm)								
	Cephamycin C			Clavulanic Acid[2]					
	#1	#2	Average	test #1	test #2	test average	control #1	control #2	control average
30	-	-	-	-	-	-	-	-	-
36	-	-	-	-	-	-	-	-	-
48	slight	slight	slight	14	11	13	-	-	-
60	12	11	12	19	16	18	-	-	-
72	12	10	11	15	17	16	-	-	-
96	21	21	21	28	28	28	11	9	10
120	25	23	24	29	28	29	16	15	16
144	25	23	24	29	28	29	17	17	17
168	27	26	27	29	28	29	21	21	21

[1] denotes zone of inhibition that was faint and diffuse, for calculations it is given a diameter of 7 mm (diameter of agar plug)

[2] The test plates detect clavulanic acid as well as the antibiotics holomycin and/or tunicamycin and/or cephamycin C, whereas the control plates detect holomycin and/or tunicamycin and/or cephamycin C. Clavulanic acid is presumed to be produced when the zone of inhibition on the test plate is larger than the zone of inhibition on the control plate at a given time.

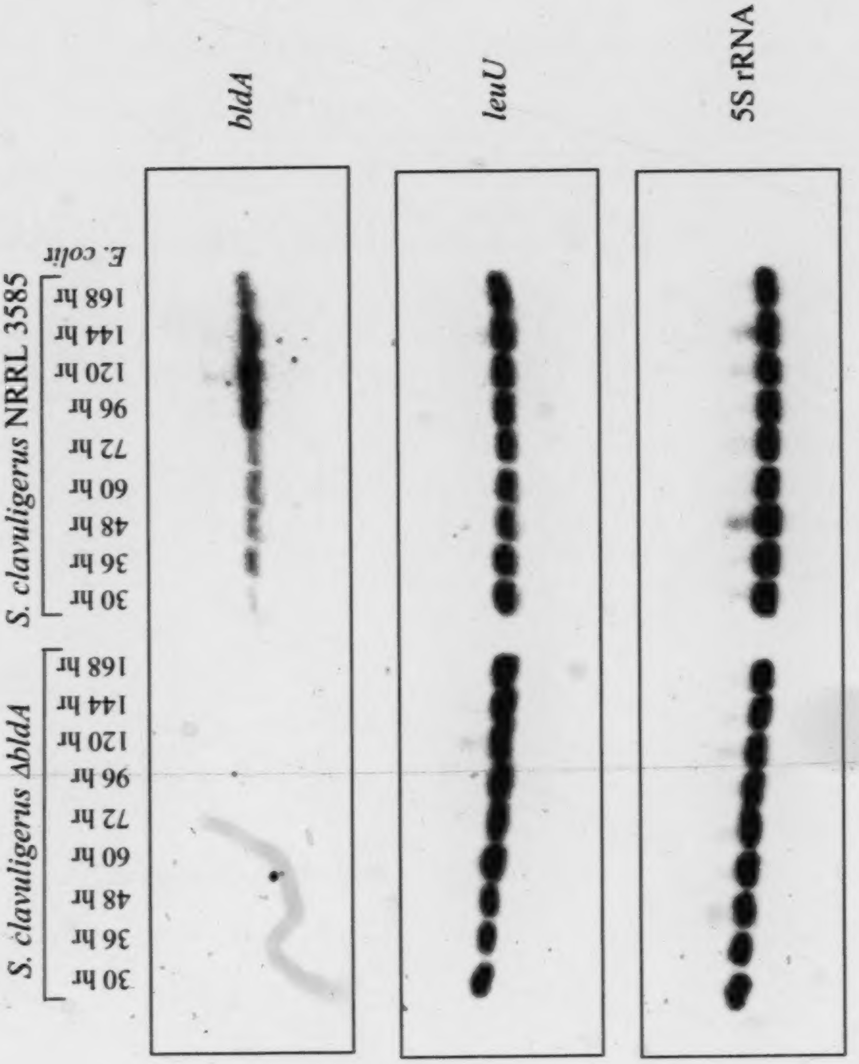
of inhibition of the clavulanic acid tests plates at this time is due in part to their presence. Maximum cephamycin C production was reached at 96 hours for the wild-type culture and 120 hours for the *bldA* culture.

RNA (5 µg) was separated on 8% denaturing polyacrylamide gels and transferred to nylon membranes. *E. coli* RNA was included as a negative control. The membranes were probed with the ^{32}P-end-labelled 21-mer, synthetic oligonucleotide BKL53 (5'-CCCTGCAGTACCATCGGCGCT-3') at 60°C to detect 5S rRNA transcripts as a control for sample loading. The membranes were then stripped and probed with the ^{32}P-end-labelled 17-mer, synthetic oligonucleotide BKL5 (5'-TTAAGCTCGCCGTGTCT-3') at 37°C to detect *bldA* transcripts, and then stripped and probed at 56°C with the ^{32}P-end-labelled oligonucleotide, BKL42 (5'-CTCAAGCTAGCGCGTCTG-3') to detect *leuU* transcripts (Figure III.9.1). Similar to the temporal pattern of accumulation observed in *S. coelicolor* J1501, the processed form of the *bldA* tRNA in *S. clavuligerus* NRRL 3585 (wild-type) is pr[illegible] low levels during vegetative growth (30 hours) and the level of tRNA increases to maximum levels at 120 hours. This is in contrast to the pattern of accumulation of the processed form of an abundant leucyl-tRNA (*leuU*), which is present at relatively constant levels throughout growth. As expected, the *bldA* tRNA is undetectable in the Δ*bldA* mutant. The *E. coli* RNA serves as a negative control for the detection of *leuU* and 5SrRNA transcripts, as the *E. coli* tRNA$^{Leu}_{GAG}$ would have three mismatches to the BKL42 probe and the *E. coli* 5SrRNA would have three mismatches to the 5S rRNA probe.

In the absence of a cognate tRNA it is expected that any genes containing the codon specified by the absent tRNA would not be fully translated. The pathway-specific regulator, *ccaR* (Perez-Llarena *et al.*, 1997), which positively regulates both

Figure III.9.1

Northern blot analysis of the *S. clavuligerus bldA* tRNA in RNA samples (5 µg) isolated at various times (hours post inoculation) from surface-grown cultures of *S. clavuligerus ΔbldA* and the isogenic parental strain *S. clavuligerus* NRRL 3585 (wild-type). The *S. clavuligerus ΔbldA* 30 and 36 hour samples are from vegetative (preantibiotic) cultures. The wild-type 30 hour sample is from vegetative cultures and the 36 hour sample is from cultures just starting to produce antibiotics. By 48 hours both wild-type and *bldA* cultures were producing both cephamycin C and clavulanic acid as determined by bioassay. The membrane was probed with the ^{32}P-end-labelled 21-mer, synthetic oligonucleotide BKL53 (5'-CCCTGCAGTACCATCGGCGCT-3') at 60°C to detect 5S rRNA transcripts as a control for sample loading. The membrane was stripped and probed with the ^{32}P-end-labelled 17-mer, synthetic oligonucleotide BKL5 (5'-TTAAGCTCGCCGTGTCT-3') at 37°C to detect *bldA* transcripts. The membrane was then stripped and probed at 56°C with the ^{32}P-end-labelled oligonucleotide, BKL42 (5'-CTCAAGCTAGCGCGTCTG-3') to detect *leuU* transcripts. *E. coli* RNA was included as a negative control.


S. clavuligerus ΔbldA
30 hr
36 hr
48 hr
60 hr
72 hr
96 hr
120 hr
144 hr
168 hr
S. clavuligerus NRRL 3585
30 hr
36 hr
48 hr
60 hr
72 hr
96 hr
120 hr
144 hr
168 hr
E. coli
bldA
leuU
5S rRNA

cephamycin C and clavulanic acid production contains a TTA codon so it would not be expected to be fully translated in the *bldA* mutant strain. It is therefore predicted that cephamycin C and clavulanic acid should not be produced in a *bldA* mutant strain. However, the Δ*bldA* mutant strain does produce both of these antibiotics which would seem to indicate that the *ccaR* gene is translated into a functional protein even in the absence of the *bldA* tRNA. If the *ccaR* transcripts are being mistranslated to produce functional proteins, then transcripts from some of the cephamycin C and clavulanic acid pathway specific genes, which require CcaR for their transcription, should be detectable by Northern analysis. Likewise, although the *ccaR* gene should be transcribed normally in the *bldA* mutant strain, there may be some difference in the level of transcripts between the wild-type and *bldA* strain. This would be particularly apparent if the *ccaR* gene is negatively autoregulated, similar to the negative autoregulation observed with another *bldA* target and pathway-specific activator, the RedZ protein of *S. coelicolor*, which regulates it's own transcription as well as the transcription of the undecylprodigiosin biosynthetic genes. Attempts were made to detect *ccaR* transcripts as well as transcripts from one of the early genes in the cephamycin C biosynthetic pathway *(pcbC)*, and one of the early genes in the clavulanic acid biosynthetic pathway *(cas2)*, using the RNA previously isolated from surface-grown cultures of *S. clavuligerus*. In all cases northern analysis failed to detect a single hybridizing band; rather a smear was detected indicating that the RNA was partially degraded and insufficient for detecting longer transcripts.

Since the temporal pattern of accumulation of the *bldA* tRNA had been previously demonstrated in surface-grown cultures, and since *S. clavuligerus* produces antibiotics in liquid culture it was decided that RNA should be isolated from liquid cultures and the pattern of expression of *bldA*, *ccaR*, *pcbC*, and *cas2* be tested under those growth conditions. RNA was therefore isolated in duplicate at 24, 36, 48, and 72 hours (post inoculation) from *S. clavuligerus* NRRL 3585 (wild-type) and *S. clavuligerus ΔbldA* liquid cultures (culture conditions described in section II.2.7). Antibiotic production was determined by bioassay (Table III.9.2). The wild-type culture produced cephamycin C at the earliest time-point although the zones were small and faint. Clavulanic acid was not detected until 72 hours in the wild-type culture and the zones of inhibition were faint and not well defined. The *bldA* cultures did not produce cephamycin C until 36 hours and clavulanic acid was not detectable, even at 72 hours. However, the bioassay results are consistent with the results from the surface-grown cultures, under the same growth conditions antibiotic production in *bldA* mutant cultures is somewhat delayed compared to wild-type cultures.

RNA was separated on a denaturing polyacrylamide gel, electroblotted and probed with end-labelled oligonucleotides specific for the *bldA* tRNA, *leuU* tRNA, and 5S rRNA transcripts as described previously (Figure III.9.2). Similar to analysis of RNA from surface-grown cultures, *bldA* transcripts are absent from the mutant strain and show a temporal pattern of accumulation in the wild-type strain. *leuU* transcripts are present at comparable levels in both strains throughout growth.

To detect the longer transcripts corresponding to *ccaR*, *cas2*, and *pcbC*, RNA (40 μg) was denatured with glyoxal and separated on a 1.25% agarose gel. *E. coli* RNA was included as a negative control for detection of *ccaR*, *pcbC*, and *cas2* transcripts, and a positive control for the detection of 16S rRNA transcripts as the 16S rRNA probe is complementary to 16S rRNA sequence from both *Streptomyces* and *E. coli*. The RNA

Table III.9.2 Antibiotic Bioassay of Duplicate Broth Cultures of *S. clavuligerus*

S. clavuligerus NRRL 3585 (wild-type)

Sample Time	Zone of inhibition (diameter in mm)								
	Cephamycin C			Clavulanic Acid					
	#1	#2	Average	test #1	test #2	test average	control #1	control #2	control average
24	15	15	15	-	-	-	-	-	-
36	28	30	29	-	-	-	-	-	-
48	35	36	36	-[1]	24	17	-	-	-
72	40	37	39	17	18	18	-	-	-

S. clavuligerus Δ*bldA* 4-1

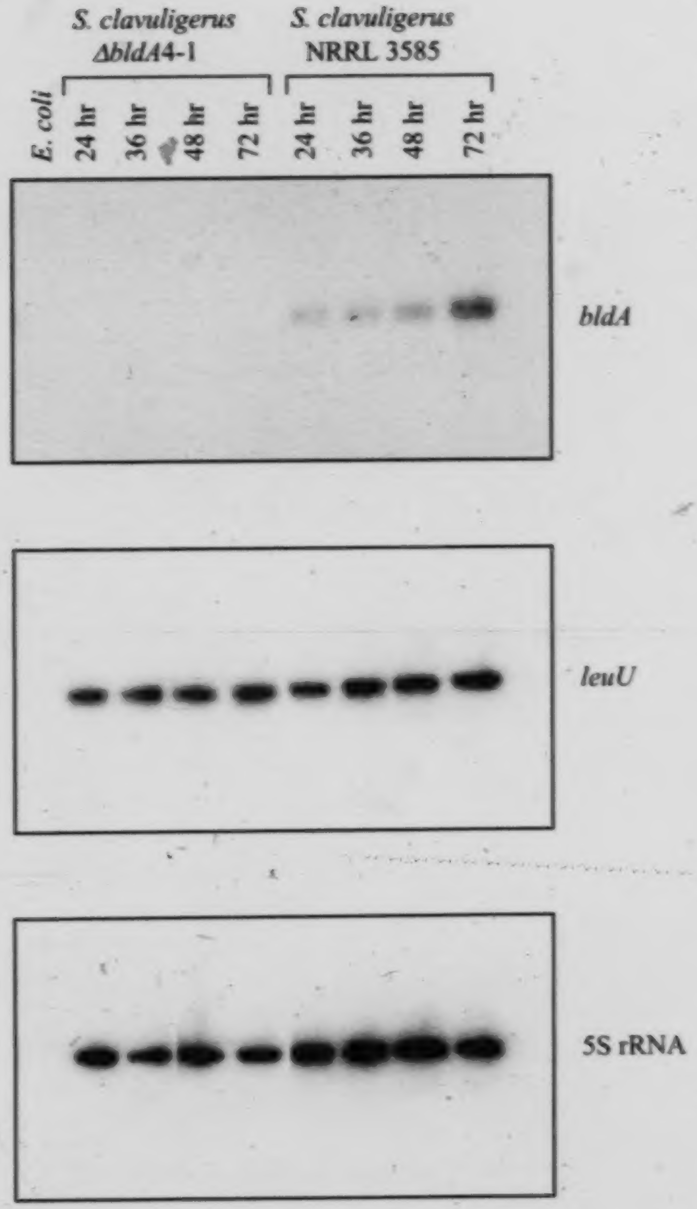
Sample Time	Zone of inhibition (diameter in mm)								
	Cephamycin C			Clavulanic Acid					
	#1	#2	Average	test #1	test #2	test average	control #1	control #2	control average
24	-	-	-	-	-	-	-	-	-
36	15	14	15	-	-	-	-	-	-
48	17	21	20	-	-	-	-	-	-
72	36	36	36	-	-	-	-	-	-

[1] For the purpose of calculating the average diameter a value of 10 mm (the diameter of the disk) was used

[2] The test plates detect clavulanic acid as well as the antibiotics holomycin and/or tunicamycin and/or cephamycin C, whereas the control plates detect only holomycin and/or tunicamycin and/or cephamycin C. Clavulanic acid is presumed to be produced when the zone of inhibition on the test plate is larger than the zone of inhibition on the control plate at a given time.

Figure III.9.2

Northern blot analysis of the *S. clavuligerus bldA* tRNA in RNA samples (5 µg) isolated at various times (hours post inoculation) from liquid cultures of *S. clavuligerus ΔbldA* and the isogenic parental strain *S. clavuligerus* NRRL 3585 (wild-type). The membrane was probed with the ^{32}P-end-labelled 17-mer, synthetic oligonucleotide BKL5 (5'-TTAAGCTCGCCGTGTCT-3') at 37°C to detect *bldA* transcripts. The membrane was then stripped and probed at 56°C with the ^{32}P-end-labelled oligonucleotide, BKL42 (5'-CTCAAGCTAGCGCGTCTG-3') to detect *leuU* transcripts. The membrane was then stripped and probed with the ^{32}P-end-labelled 21-mer, synthetic oligonucleotide BKL53 (5'-CCCTGCAGTACCATCGGCGCT-3') at 60°C to detect 5S rRNA transcripts to control for sample loading. *E. coli* RNA was included as a negative control.


S. clavuligerus
ΔbldA4-1
S. clavuligerus
NRRL 3585
E. coli
24 hr
36 hr
48 hr
72 hr
24 hr
36 hr
48 hr
72 hr
bldA
leuU
5S rRNA

was transferred to a nylon membrane and probed with a ^{32}P-dCTP random primer-labelled 574 bp *Nco*I-*Sac*I *ccaR* gene fragment from the plasmid pWOR2925 (Figure III.9.3). The plasmid pWOR2925 containing the *ccaR* gene contains an *Nco*I site engineered at the *ccaR* start codon and has the TTA codon replaced with a CTA codon at amino acid 32 (D. Alexander, personal communication). These changes generate two mismatches between the probe sequence and the *ccaR* transcript, which should not be problematic with a 574 bp probe. Hybridization was carried out at 50°C in 50% formamide hybridization solution. Washes were carried out at 50°C, with a final wash at 60°C. The membrane was stripped and probed with the ^{32}P-end-labelled 18-mer oligonucleotide probe A (5'-ACTATCGGAGAGGCCATG-3')(Paradkar and Jensen, 1995) to detect *cas2* transcripts (Figure III.9.3). Probe A is complementary to $^{-}$1-$^{+}$17 of the *cas2* transcript. Hybridization and washes were carried out at 45°C. The membrane was stripped and probed with ^{32}P-end-labelled BKL54 (5'-CCGCCTTCGCCACCGGT-3') to detect 16S rRNA transcripts as a control for sample loading. A duplicate membrane was probed with a 0.7 kb fragment complementary to the *pcbC* transcript (Figure III.9.3). The *pcbC* probe was purified as an *Nde*I-*Kpn*I fragment from the plasmid pMDW (Durairaj *et al.*, 1996) and random primer-labelled with ^{32}P-dCTP. Hybridization was carried out at 45°C in hybridization solution containing 50% formamide. Washes were carried out at 45°C. The results of the northern analysis are shown in Figure III.9.4. As expected, *ccaR* transcripts are present in the *ΔbldA* mutant. The levels of transcript are lower than in the wild-type strain, which could be the reason why antibiotic production in the *bldA* mutant is somewhat delayed. If CcaR was a repressor of its own synthesis, higher levels of transcript would be expected if the translation rate of *ccaR* transcripts was impeded in the

Figure III.9.3

A) Schematic diagram of the cephamycin C and clavulanic acid gene super cluster. The two genes clusters are contiguous with the clavulanic acid biosynthetic genes immediately downstream of the *pcbR* gene as indicated. **B)** The gene fragments and oligonucleotide (probe A) that were used to probe northern membranes are indicated.

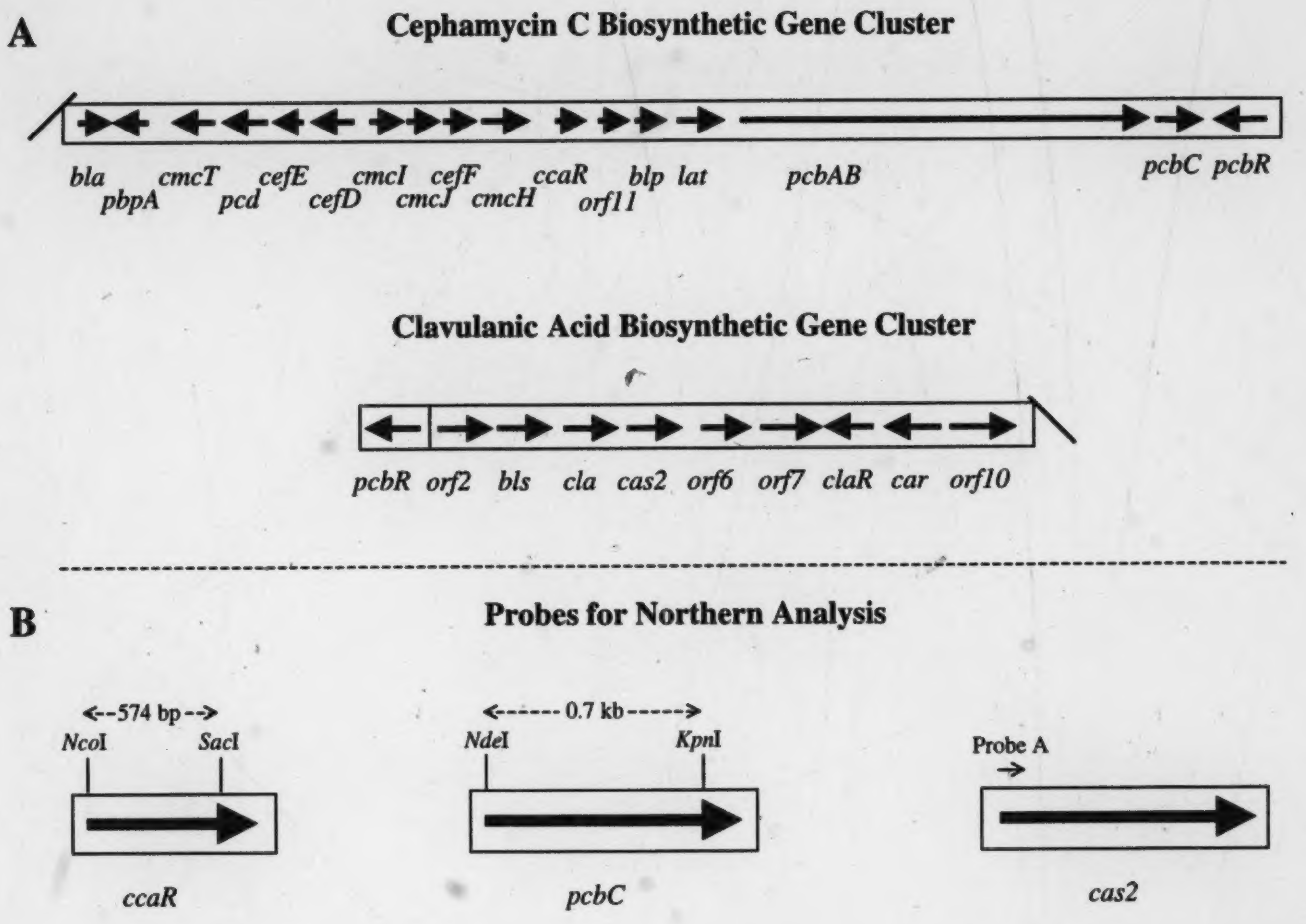
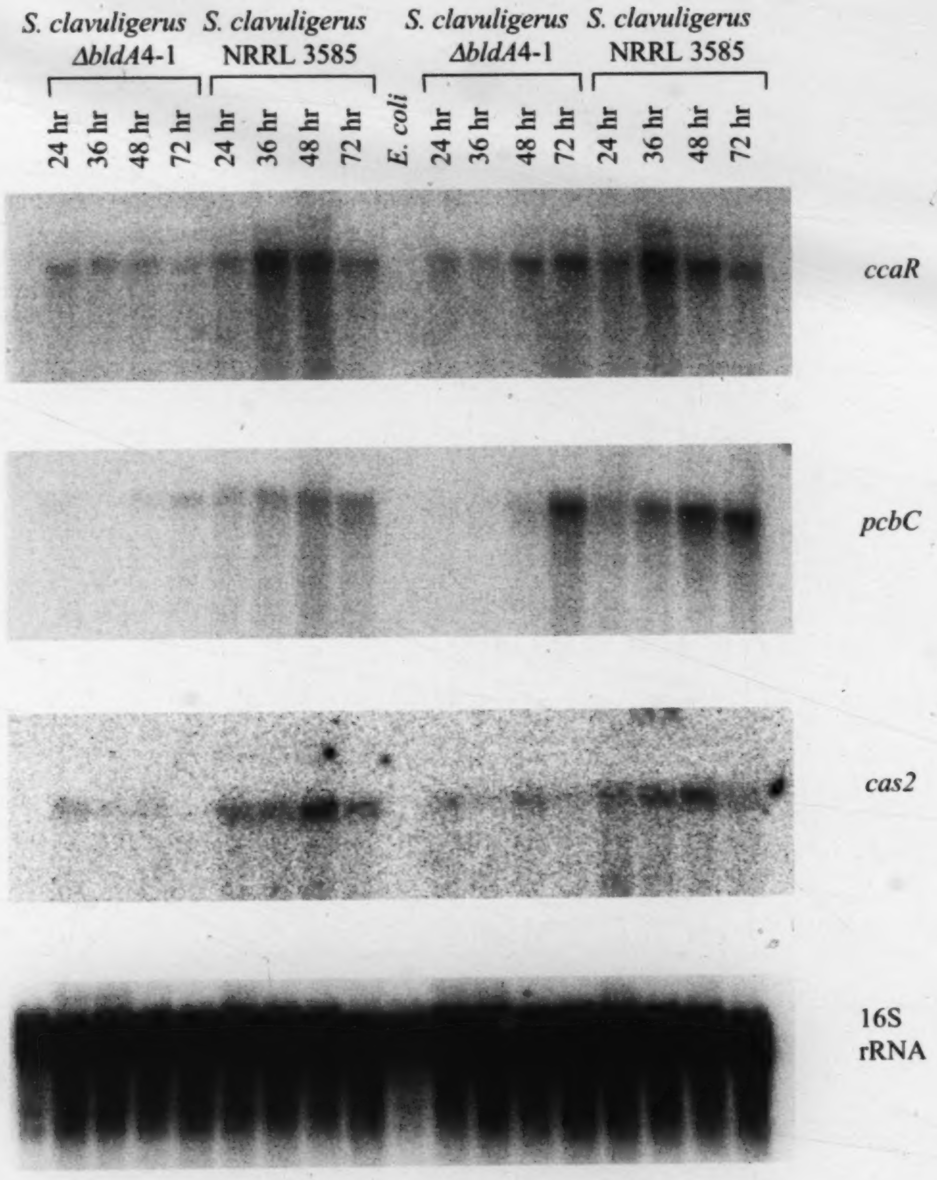

A
Cephamycin C Biosynthetic Gene Cluster
bla
pbpA
cmcT
pcd
cefE
cefD
cmcI
cmcJ
cefF
cmcH
ccaR
orf11
blp
lat
pcbAB
pcbC
pcbR
Clavulanic Acid Biosynthetic Gene Cluster
pcbR
orf2
bls
cla
cas2
orf6
orf7
claR
car
orf10
B
Probes for Northern Analysis
<--574 bp-->
NcoI
SacI
ccaR
<------ 0.7 kb ------>
NdeI
KpnI
pcbC
Probe A
cas2

Figure III.9.4

Northern blot analysis of the *S. clavuligerus* *ccaR, pcbC,* and *cas2* transcripts in RNA samples isolated at various times (hours post inoculation) from liquid cultures of *S. clavuligerus ΔbldA* and the isogenic parental strain *S. clavuligerus* NRRL 3585 (wild-type). RNA (40 µg) was denatured with glyoxal and separated on a 1.25% agarose gel. *E. coli* RNA was included as a negative control for the detection of *ccaR, pcbC,* and *cas2* transcripts and as a positive control for the detection of 16S rRNA transcripts. The RNA was transferred to a nylon membrane and probed with a ^{32}P-dCTP random primer-labelled 574 bp *Nco*I-*Sac*I fragment from the *ccaR* gene. Hybridization was carried out at 50°C in 50% formamide hybridization solution. Washes were carried out at 50°C, with a final wash at 60°C. The membrane was stripped and probed with ^{32}P-end-labelled 18-mer oligonucleotide probe A (5'-ACTATCGGAGAGGCCATG-3') to detect *cas2* transcripts. Hybridization and washes were carried out at 45°C. The membrane was stripped and probed with ^{32}P-end-labelled BKL54 (5'-CCGCCTTCGCCACCGGT-3') to detect 16S rRNA transcripts to control for RNA loading. A duplicate membrane was probed with a ^{32}P-dCTP random primer-labelled 0.7 kb fragment complementary to the *pcbC* transcript. Hybridization was carried out at 45°C in hybridization solution containing 50% formamide. Washes were carried out at 45°C.


S. clavuligerus ΔbldA4-1
S. clavuligerus NRRL 3585
S. clavuligerus ΔbldA4-1
S. clavuligerus NRRL 3585
24 hr
36 hr
48 hr
72 hr
24 hr
36 hr
48 hr
72 hr
E. coli
24 hr
36 hr
48 hr
72 hr
24 hr
36 hr
48 hr
72 hr
ccaR
pcbC
cas2
16S rRNA

bldA strain. Since there are fewer *ccaR* transcripts detected, it appears that *ccaR* is not negatively autoregulated. Although there appears to be lower levels of *ccaR* transcript in the *bldA* strain as compared to the wild-type strain, this could be a function of inoculum density as the timing and amount of cephamycin C and clavulanic acid produced in *S. clavuligerus* is affected by the cell density of the inoculum (Sanchez and Brana, 1996). Although efforts were made to inoculate both wild-type and *bldA* mutant cultures to the same OD_{600}, the actual number of colony forming units could vary significantly, and as the wild-type and mutant strain have different morphologies it will be difficult to make meaningful comparisons. *pcbC* and *cas2* transcripts are also detected in the mutant strain which indicates that the *ccaR* transcripts are being translated into functional proteins, however, the level of *pcbC* and *cas2* transcripts are clearly lower in the *bldA* mutant than in the wildtype strain. This likely indicates that less functional CcaR protein is produced in the *bldA* strain, either due to the overall lower levels of *ccaR* transcript available to be translated, or due to an increase in the amount of: aborted translation, or mistranslation of *ccaR* transcripts. Although clavulanic acid was not detected in the *bldA* mutant culture supernatant from this particular liquid-culture time course, the presence of *cas2* transcripts indicates that clavulanic acid pathway is activated. Since clavulanic acid was not detected in the wild-type culture supernatant until 72 hours and antibiotic production in the mutant strain is somewhat delayed it is probable that clavulanic acid would have been detected at later time points.

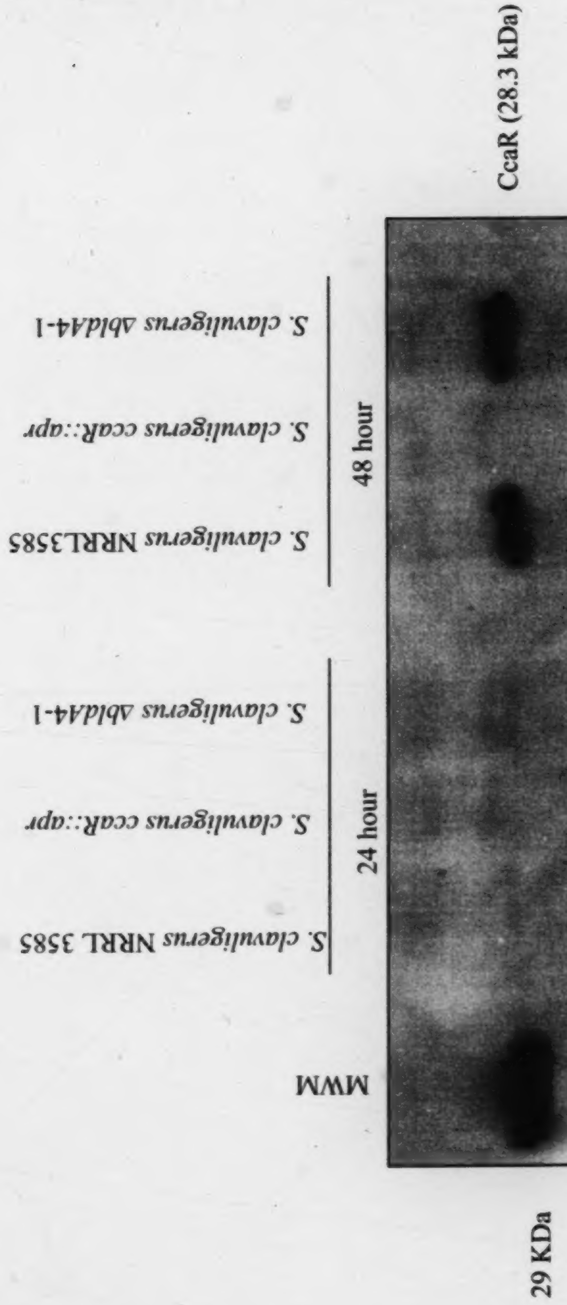
Although the presence of *pcbC* and *cas2* transcripts, as well as the detection of cephamycin C and clavulanic acid, provides indirect evidence that the *ccaR* transcripts are translated in the *bldA* mutant, it was of interest to determine if the level of CcaR protein was consistent with the amount of transcript to provide some indication of the efficiency of translation of the single UUA codon. Cell free extracts were obtained from *S. clavuligerus ΔbldA* and *S. clavuligerus* NRRL 3585 (wild-type) cultures at 24 and 48 hours. Cell free extracts from the *S. clavuligerus ccaR::apr* disruption strain (Alexander and Jensen, 1998) were used as a negative control. The proteins were separated by SDS-PAGE gel electrophoresis, transferred to a PVDF membrane and probed with α-CcaR antibodies (Figure III.9.5). At 24 hours low levels of CcaR protein are present in both the wild-type and *ΔbldA* samples, and both cultures show considerable amounts of CcaR after 48 hours of growth. There appears to be no noticeable difference in amount of CcaR protein between the wild-type and *bldA* mutant cultures which indicates that *ccaR* transcripts are efficiently translated in the *bldA* strain. The comparable levels of CcaR protein in the mutant and wild-type strain differ somewhat from the RNA analysis where lower levels of *ccaR, cas2,* and *pcbC* transcripts were detected in the *bldA* mutant. The difference between level of *ccaR* transcript and CcaR protein could indicate that the *ccaR* transcripts in the *bldA* mutant are more stable, however, since the RNA and protein samples were not isolated from the same set of cultures it could also reflect differences in growth of the particular cultures or inoculum density.

III.10 REPORTER GENE ANALYSIS

At the time of study the only known *bldA* target (TTA codon containing gene) in *S. clavuligerus* was *ccaR*. While both the northern and western analysis clearly confirmed that *ccaR* transcripts are translated into functional protein in the absence of the *bldA* tRNA, it was unknown if the apparently high levels of mistranslation are a

Figure III. 9.5

Western analysis of the *S. clavuligerus* CcaR protein in cell free extracts from *S. clavuligerus* NRRL 3585 (wild-type), *S. clavuligerus ΔbldA*4-1, and *S. clavuligerus ccaR::apr* isolated after 24 and 48 hours of growth in TBS +1% starch broth. Samples (10 μg) were separated on a 10% SDS-PAGE gel, transferred to a PVDF membrane and reacted with rabbit α-CcaR antibodies. Goat α-rabbit immunoglobulin G horseradish peroxidase conjugate was used as the secondary antibody and the protein was detected using the ECL (enhanced chemiluminescent) western system reagents and protocol.


29 KDa
MWM
S. clavuligerus NRRL 3585
S. clavuligerus ccaR::apr
S. clavuligerus ΔbldA4-1
24 hour
S. clavuligerus NRRL3585
S. clavuligerus ccaR::apr
S. clavuligerus ΔbldA4-1
48 hour
CcaR (28.3 kDa)

characteristic of *S. clavuligerus* or if the effect is more context dependant. In order to address this issue, the translation efficiency of a TTA codon-containing reporter gene was analyzed. The reporter gene chosen for this study was the *mut3* version of the gene for green fluorescent protein (GFP) (Cormack *et al.*, 1996). This version of GFP fluoresces 100× more brightly than the wild-type protein from the jellyfish *Aequorea victoria,* and contains three TTA codons. The sequence of *gfp-mut3* is shown in Figure III.10.1. GFP was chosen over some of the conventional *Streptomyces* reporter genes for several reasons. It doesn't require selection for expression, it is not toxic to *Streptomyces*, it is easy to detect and can be monitored throughout growth. There was some initial concern that the *gfp-mut3* gene might be expressed poorly in *Streptomyces* as it contains many codons that are rarely used in the high G+C *Streptomyces* mRNA (54.6% of the *gfp-mut3* codons are used at a frequency of <5/1000 codons in *Streptomyces* genes), however, *S. lividans* colonies containing the *gfp-mut3* gene on a high copy number vector fluoresced as brightly as colonies containing the *hgfp* gene which has been optimized for expression in human cell lines and contains primarily abundant *Streptomyces* codons (Markus and Leskiw, 1997).

The *gfp-mut3* gene under the control of the constitutive *ermE** promoter (Bibb *et al.*, 1985; Bibb *et al.*, 1986) on the high copy number *Streptomyces* plasmid pIJ486 was generously provided by H. Markus (University of Alberta). The plasmid was transformed into *S. clavuligerus* NRRL 3585 (wild-type) and into *S. clavuligerus* Δ*bldA*4-1. As expected, the wild-type transformants expressed GFP and fluoresced under a mercury lamp, while the *bldA* transformants did not express GFP (Figure III.10.2). To verify that the differences in expression of GFP between the two strains was due to the

Figure III.10.1

Sequence of the *mut3gfp* gene. The ribosome binding site is indicated by a line and the three TTA codons are highlighted in bold text. Amino acids indicated by bold underlined text translate rare codons (codons that are used at a frequency of <1/1000 codons in *Streptomyces* genes). Amino acids indicated by bold text translate minor codons that are used at a frequency of 1-5/1000 codons. Adapted from the table of codon usage for the genus *Streptomyces* obtained from the Online Library at the Streptomyces Internet Resource Center http://molbio.cbs.umn.edu/asirc/lib/lib.html

```
         RBS            M  S  K  G  E  E  L  F  T  G  V  V
TAGATTTAAGAAGGAGATATACATATGAGTAAAGGAGAAGAACTTTTCACTGGAGTTGTC    60
ATCTAAATTCTTCCTCTATATGTATACTCATTTCCTCTTCTTGAAAAGTGACCTCAACAG

P  I  L  V  E  L  D  G  D  V  N  G  H  K  F  S  V  S  G  E
CCAATTCTTGTTGAATTAGATGGTGATGTTAATGGGCACAAATTTTCTGTCAGTGGAGAG   120
GGTTAAGAACAACTTAATCTACCACTACAATTACCCGTGTTTAAAAGACAGTCACCTCTC

G  E  G  D  A  T  Y  G  K  L  T  L  K  F  I  C  T  T  G  K
GGTGAAGGTGATGCAACATACGGAAAACTTACCCTTAAATTTATTTGCACTACTGGAAAA   180
CCACTTCCACTACGTTGTATGCCTTTTGAATGGGAATTTAAATAAACGTGATGACCTTTT

L  P  V  P  W  P  T  L  V  T  T  F  S  Y  G  V  Q  C  F  S
CTACCTGTTCCATGGCCAACACTTGTCACTACTTTCTCTTATGGTGTTCAATGCTTTTCA   240
GATGGACAAGGTACCGGTTGTGAACAGTGATGAAAGAGAATACCACAAGTTACGAAAAGT

R  Y  P  D  H  M  K  Q  H  D  F  F  K  S  A  M  P  E  G  Y
AGATACCCAGATCATATGAAACAGCATGACTTTTTCAAGAGTGCCATGCCCGAAGGTTAT   300
TCTATGGGTCTAGTATACTTTGTCGTACTGAAAAAGTTCTCACGGTACGGGCTTCCAATA

V  Q  E  R  T  I  F  F  K  D  D  G  N  Y  K  T  R  A  E  V
GTACAGGAAAGAACTATATTTTTCAAAGATGACGGGAACTACAAGACACGTGCTGAAGTC   360
CATGTCCTTTCTTGATATAAAAAGTTTCTACTGCCCTTGATGTTCTGTGCACGACTTCAG

K  F  E  G  D  T  L  V  N  R  I  E  L  K  G  I  D  F  K  E
AAGTTTGAAGGTGATACCCTTGTTAATAGAATCGAGTTAAAAGGTATTGATTTTAAAGAA   420
TTCAAACTTCCACTATGGGAACAATTATCTTAGCTCAATTTTCCATAACTAAAATTTCTT

D  G  N  I  L  G  H  K  L  E  Y  N  Y  N  S  H  N  V  Y  I
GATGGAAACATTCTTGGACACAAATTGGAATACAACTATAACTCACACAATGTATACATC   480
CTACCTTTGTAAGAACCTGTGTTTAACCTTATGTTGATATTGAGTGTGTTACATATGTAG

M  A  D  K  Q  K  N  G  I  K  V  N  F  K  I  R  H  N  I  E
ATGGCAGACAAACAAAAGAATGGAATCAAAGTTAACTTCAAAATTAGACACAACATTGAA   540
TACCGTCTGTTTGTTTTCTTACCTTAGTTTCAATTGAAGTTTTAATCTGTGTTGTAACTT

D  G  S  V  Q  L  A  D  H  Y  Q  Q  N  T  P  I  G  D  G  P
GATGGAAGCGTTCAACTAGCAGACCATTATCAACAAAATACTCCAATTGGCGATGGCCCT   600
CTACCTTCGCAAGTTGATCGTCTGGTAATAGTTGTTTTATGAGGTTAACCGCTACCGGGA

V  L  L  P  D  N  H  Y  L  S  T  Q  S  A  L  S  K  D  P  N
GTCCTTTTACCAGACAACCATTACCTGTCCACACAATCTGCCCTTTCGAAAGATCCCAAC   660
CAGGAAAATGGTCTGTTGGTAATGGACAGGTGTGTTAGACGGGAAAGCTTTCTAGGGTTG

E  K  R  D  H  M  V  L  L  E  F  V  T  A  A  G  I  T  H  G
GAAAAGAGAGACCACATGGTCCTTCTTGAGTTTGTAACAGCTGCTGGGATTACACATGGC   720
CTTTTCTCTCTGGTGTACCAGGAAGAACTCAAACATTGTCGACGACCCTAATGTGTACCG

M  D  E  L  Y  K  *
ATGGATGAACTATACAAATAAATGTCCAGACCTGCAG        757
TACCTACTTGATATGTTTATTTACAGGTCTGGACGTC
```

Figure III.10.2

S. clavuligerus NRRL 3585 colony transformed with the *mut3gfp* recombinant plasmid **(A)** 10× magnification, 1 second exposure under normal illumination **(B)** 10× magnification, 3 second exposure under a mercury lamp with FITC filter. *S. clavuligerus* Δ*bldA*4-1 colony transformed with the *mut3gfp* recombinant plasmid **(C)** 10× magnification, 1 second exposure under normal illumination **(D)** 10× magnification, 3 second exposure under a mercury lamp with FITC filter.

A
B
C
D

inability of the *bldA* strain to translate the TTA codons in the *gfp* transcript, and not due to the result of rearrangement or deletion of the *gfp* construct in the *bldA* strain, plasmid DNA was isolated from both wild-type and Δ*bldA* transformants and back-transformed into the corresponding isogenic Δ*bldA* and wild-type strains. When reintroduced into the wild-type strain, plasmid from either wild-type or Δ*bldA* transformants expressed GFP. In contrast, plasmids reintroduced into the Δ*bldA* strain did not express GFP regardless of the host from which they were isolated. The results demonstrate that not all TTA codons are mistranslated efficiently in the *S. clavuligerus bldA* mutant strain.

III.11 *S1* NUCLEASE PROTECTION ANALYSIS

In order to identify the promoter and transcription start site of the *S. clavuligerus bldA* gene, *S1* nuclease protection analysis was performed. The promoter region of the *S. coelicolor bldA* gene shows significant homology to the sequence upstream from the *S. clavuligerus bldA* tRNA gene (Figure III.11.1) so the transcription start site was predicted to be at -70 or -69. The probe for *S1* nuclease mapping of the *bldA* gene was generated by PCR amplification of a 232 bp fragment with p9S+R as template DNA. The primers were an 18-mer synthetic oligonucleotide NTR10 (5'-CGGAGCCGGACTCGAACC-3') corresponding to the sequence near the 3' end of the tRNA gene and a 27-mer synthetic oligonucleotide NTR9 (5'-CCGCGAATTCGCCATGGAACGCCTTGT-3') corresponding to the region 128-147 bases upstream of the tDNA and 60 bases upstream from the expected transcription start point (see Figure III.11.2). The upstream primer contains a 7 nucleotide non-homologous extension (underlined) at the 5' end which negates the need to separate the strands after probe labelling since the 5' labelled non-homologous end will be removed by *S1* nuclease

Figure III.11.1

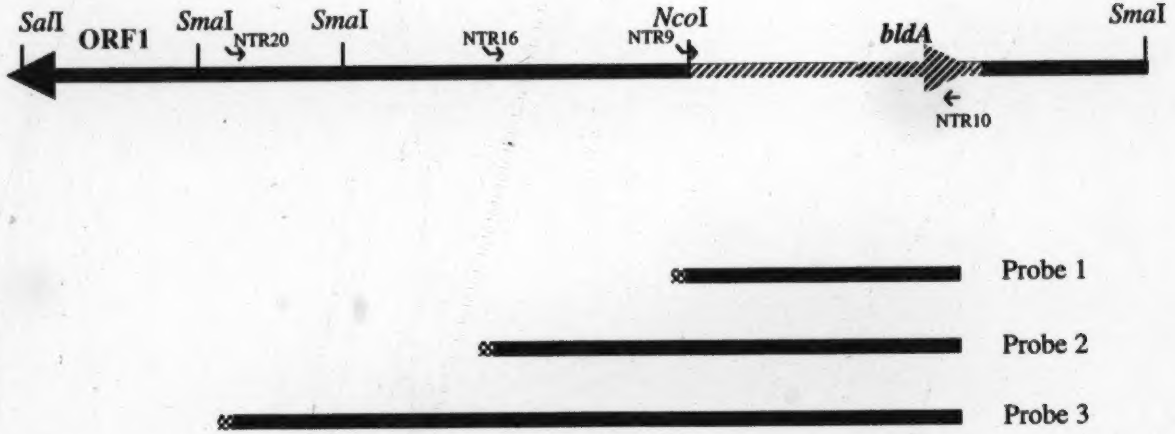
Alignment of the *S. coelicolor bldA* and *S. clavuligerus bldA* tRNA genes and promoter regions. The –35 and –10 promoter region of the *S. coelicolor bldA* gene is indicated as well as the transcription start point (corresponds to one of two G residues marked by ⇨ , tsp). The start of the tDNA is also indicated for both sequences.

```
                  -35                    -10           ↓↓ tsp
S. coelicolor   - CTGATGCACCCTCAGGCGCCAATAAAATAAGACAACATCGGCCGTCCCGC    50
                  |||| ||||||||| ||| |||||||||||||||||||||| |  |  ||
S. clavuligerus - CTGACGCACCCTCAAGCGGCAATAAAATAAGACAACATCGGGCCACGGGC    50

S. coelicolor   - CCGAGGCCGGGAGGAGCGTCCCG-ATAGGGTCAAGGCGCGCGAAGCGGAG    99
                  |   || || |  || |  |||| ||||| ||| |  ||   | ||||||
S. clavuligerus - CGCCGGGCGAGTCGATCTCCCCGGATAGGCTCATGT-GCTT-ACGCGGAG    98

                  tRNA start ↘
S. coelicolor   - CCCATGGCCTG-CCCGGATGGTGGAATGCAGACACGGCGAGCTTAAACCT   148
                  ||     |||| ||||||||||||||||||||||||||||||||||||||
S. clavuligerus - CCACGCTCCTGGCCCGGATGGTGGAATGCAGACACGGCGAGCTTAAACCT   148
                  tRNA start ↗

S. coelicolor   - CGCTGCCCCTTCGAGGGCGTGCCGGTTCAAGTCCGGCTCCGGGCACCA     196
                  ||||| || |||  || ||||||||||| |||||||||||||||||
S. clavuligerus - CGCTGGCC-TTCATGGCCGTGCCGGTTCGAGTCCGGCTCCGGGCAC       193
```

Figure III.11.2

Probes for *S1* nuclease protection analysis. The primers used to amplify the probes are indicated by arrows. The upstream primers contain non-homologous extensions which are removed by *S1* nuclease treatment. These short extensions are indicated by hatched boxes on the probes. The region deleted from the chromosome of the *bldA* mutant, from the *Nco*I site to the 3' end of the tRNA gene, is indicated by striped boxes. Originally probe 3 was believed to extend into the N-terminus of ORF1; it was later determined that the open reading frame lies outside of this region.

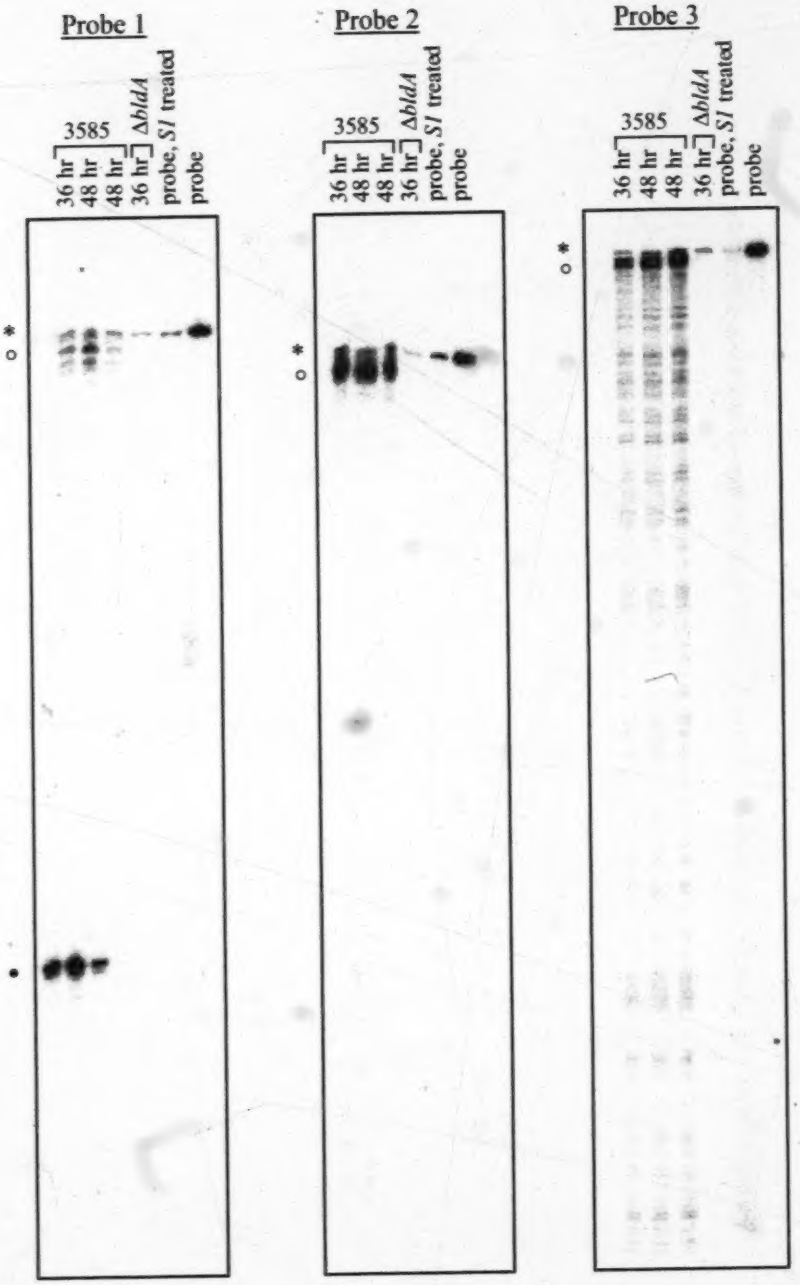
Probes for S1 Nuclease Protection Analysis


SalI
ORF1
SmaI
NTR20
SmaI
NTR16
NTR9
NcoI
bldA
SmaI
NTR10
Probe 1
Probe 2
Probe 3

treatment. The use of the non-homologous extension allows differentiation between full-length protection and probe-probe reannealing since the full- length protected fragments will be 7 nucleotides shorter than the reannealed probe fragments. RNA samples (10-40 μg) from *S. clavuligerus* NRRL 3585 (wild-type) and *S. clavuligerus Δbld*4-1 surface-grown cultures were hybridized to the ^{32}P end-labelled probe and treated with *S1* nuclease as described in section II.8.7. A sample of probe alone was treated with *S1* nuclease as a control. The results are shown in Figure III.11.3. A major band which corresponded to the processed tRNA was identified in the *S. clavuligerus* NRRL 3585 (wild-type) samples. In addition, a band which corresponded to full-length protection was present in the wild-type samples, but no bands corresponding to the predicted start of the primary transcript were observed. The experiment was repeated several times, the annealing temperature was lowered by 5° to increase the sensitivity, and higher quality RNA from broth cultures was also used, however, none of these experiments identified a primary transcript initiating at a site analogous to that of the *S. coelicolor bldA* gene. Since that primary transcript was not observed and full-length protected fragments were, additional probes were made to identify a promoter further upstream from the *bldA* tDNA (Figure III.11.2). Probe 2 and probe 3 were generated by PCR amplification using p9S+R as template DNA. The primers for probe 2 (421 bp) were NTR10 and NTR16, a 28-mer synthetic oligonucleotide (5'-GCGAGAATTCCACTCTTCGCTGCGGTGT-3') which corresponds to the region 316-333 bases upstream from the tDNA. The primers for probe 3 (593 bp) were NTR10 and NTR20, a 28-mer synthetic oligonucleotide (5'-GCTAGAATTCGGAGGCCGTCTCCGAGGA-3') which corresponds to the region 489-56 bases upstream from the tDNA. Originally, this primer was thought to correspond to the

Figure III.11.3

S1 nuclease protection analysis of *S. clavuligerus bldA* transcripts. RNA from *S. clavuligerus* NRRL 3585 (wild-type) and *S. clavuligerus ΔbldA*4-1 was annealed to ^{32}P end-labelled probes, treated with *S1* nuclease and separated on a 6% denaturing polyacrylamide gel. The RNA samples were isolated from wild-type liquid cultures 36, and 48 hours post-inoculum and from the *bldA* culture 36 hours post-inoculum. Ten micrograms of RNA was used for one set of wild-type 48 hour reactions (lane 2, all three panels), twenty micrograms was used for all other reactions. The two lanes in each panel labelled probe represent probe that has been treated with *S1* nuclease (first lane) and 3000 cpm of purified probe (second lane). The gel in the first panel (probe 1) was electrophoresed for 2 hours, the gels in the second and third panels (probe 2 and probe 3) were electrophoresed for 4 hours. Probe-probe reannealed bands are indicated by (*), full-length protected fragments are indicated by (○), and the 5' end of the processed, mature tRNA is indicated by (●).


Probe 1
Probe 2
Probe 3
3585
36 hr
48 hr
48 hr
36 hr
ΔbldA
probe, S1 treated
probe
*
o
•

N-terminus of ORF1, however, it was later determined to lie 55 bases upstream from the ORF1 start codon. As was the case with NTR9, NTR16 and NTR20 contain 5' non-homologous extensions (underlined). Probe 2 and 3 showed essentially the same results as probe 1. Full-length protected fragments were observed in addition to the band corresponding to the processed tRNA, but a primary transcript initiating within the probe sequences was not observed. Although the sequencing ladders were of poor quality and are not shown, sufficient sequence data was obtained to verify the approximate size of the bands. The reactions using probe 2 and probe 3 were electrophoresed for 4 hours to obtain better separation of the longer fragments and do not show the processed form of the tRNA. However, these bands were observed when the reactions were electrophoresed for shorter periods of time (data not shown).

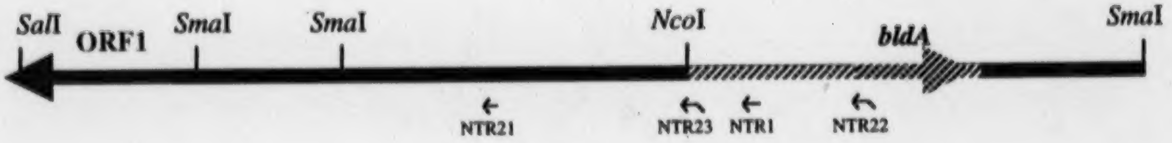
III.12 PRIMER EXTENSION ANALYSIS

Since *S1* nuclease protection analysis failed to identify the transcription start point of the *bldA* tRNA, primer extension analysis was performed. The primers used for primer extension analysis were: NTR22 (5'-ACGTCTAGACCACCATCCGGGCCAGGA-3'), a 27-mer complementary to the 5' end of the *bldA* tRNA; NTR1 (5'-TGCGTCAGCCCAGGTGG-3'), a 17-mer complementary to the region 102-119 bases upstream from the tDNA; NTR23 (5'-CGCTCTAGAAGGCGTTCCATGGCGAA-3'), a 26-mer complementary to the region 130-147 bases upstream from the tDNA; and NTR21 (5'-ACACCGCAGCGAAGAGTGA-3'), a 19-mer complementary to the region 316-335 bases upstream from the tDNA (See Figure III.12.1). NTR22 and NTR1 are homologous to sequences in the region that was deleted when the *bldA* deletion strain was generated, and should not hybridize to RNA from the Δ*bldA* strain. The NTR23 primer spans the *Nco*I

Figure III.12.1

Primers for primer extension analysis. The direction and approximate region that the primer is complementary to is indicated. The region deleted from the chromosome of the *bldA* mutant, from the *Nco*I site to the 3' end of the tRNA gene, is indicated by striped boxes. The tDNA is indicated by a black striped arrow and the partial open reading frame upstream and divergent to the *bldA* tRNA is indicated in grey.

Primers for Primer Extension Analysis


*Sal*I
ORF1
*Sma*I
*Sma*I
*Nco*I
bldA
*Sma*I
NTR21
NTR23
NTR1
NTR22

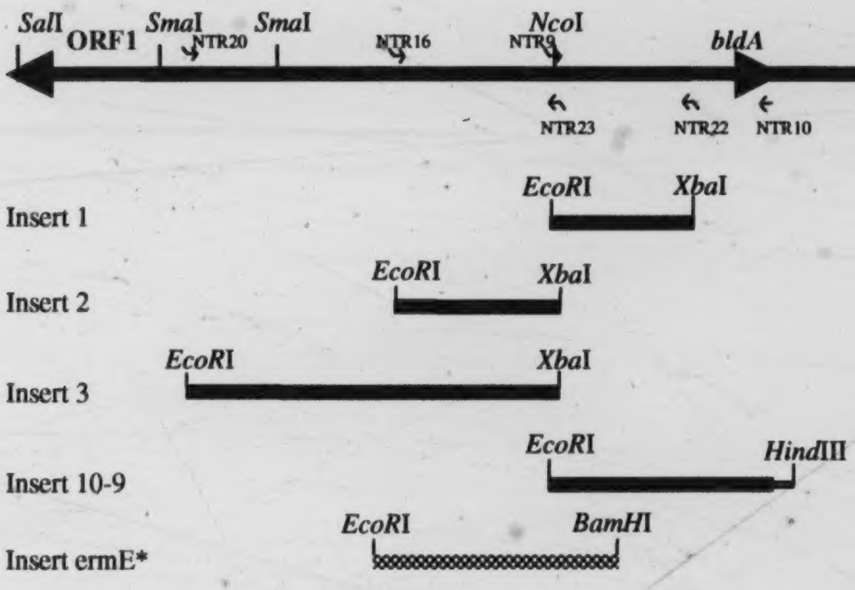
should not hybridize to RNA from the Δ*bldA* strain. The NTR23 primer spans the *Nco*I site that marks the boundary of the deleted sequence and should have only 10 bases of homology to *bldA* transcripts from the Δ*bldA* strain, which should be insufficient for stable hybrid formation. NTR21 is homologous to sequence outside of the deleted region and should yield similar results for both wild-type and mutant RNA samples. However, all of the primers yielded identical products for both wild-type and mutant samples, indicating that the primers were binding either non-specifically to RNA or binding specifically to other transcripts. Hybridization was originally carried out at 37°C, which is well below the T_d of all of the primers. The hybridization temperature was increased to 42°C, 47°C, 52°C, 56°C and 62°C without a detectable change in the results. The highest annealing temperature used (62°C) is above the T_d of NTR21 (T_d 60°C) and NTR23 (T_d 56°C) and at the T_d of NTR22 so it seems most probable that all of the primers are hybridizing to transcripts other than *bldA*. NTR22 is homologous to the 5' end of the *bldA* tDNA so it is likely hybridizing to another leucyl tDNA species. NTR23 hybridizes to a region that contains several oligonucleotide sequences that are typically found in *Streptomyces* promoters (Bourn and Babb, 1995), so it could be hybridizing to a transcript from a gene with more than one promoter. The primers NTR23 and NTR22 have non-homologous extensions at the 5' end as they were originally designed as primers for PCR cloning. The extensions contain a *Xba*I site which is relatively rare in *Streptomyces* DNA, so it seems unlikely that the extensions are causing the mispriming. Also these primers were used successfully as sequencing primers and there was no apparent difference in the results when the primer NTR1, without a non-homologous 5' extension was used. At any rate, the *bldA* promoters and transcription start points will have to be verified by other means.

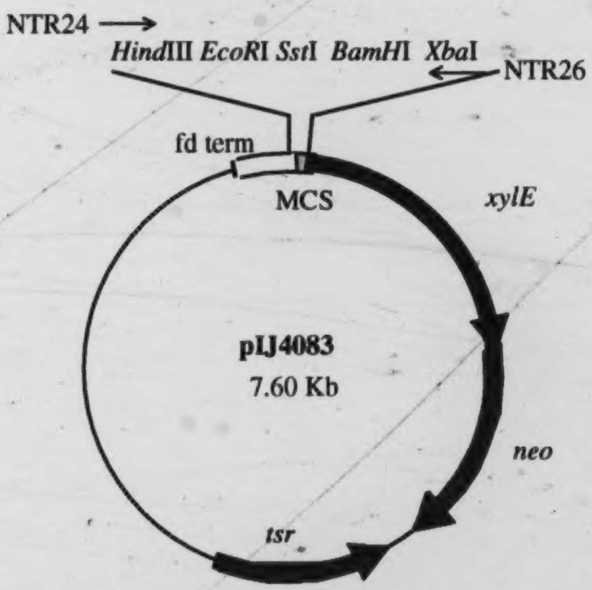
III.13 PROMOTER PROBE ANALYSIS

Since *in vitro* analysis of RNA failed to identify the *bldA* promoter and transcription start site, attempts were made to identify the promoter by *in vivo* methods. Promoter probe analysis was carried out using the reporter gene *xylE* which encodes catechol-2,3- dioxygenase (C23O). The *Streptomyces* vector pIJ4083 contains the promoter-less *xylE* gene downstream of the fd transcriptional terminator and a polylinker cloning site (Clayton and Bibb, 1990). Insertion of fragments of interest into the multiple cloning site allows the identification of active promoter sequences by comparing levels of C23O activity between transformed cultures. Fragments of interest were generated by PCR amplification using p9S+R as template DNA (See Figure III.13.1). The pairs of primers used to amplify insert 1, 2, and 3 were NTR9
(5'-CCGCGAATTCGCCATGGAACGCCTTGT-3') and
NTR22 (5'-ACGTCTAGACCACCATCCGGGCCAGGA-3'),
NTR16 (5'-GCGAGAATTCCACTCTTCGCTGCGGTGT-3') and
NTR23 (5'-CGCTCTAGAAGGCGTTCCATGGCGAA-3'), and
NTR20 (5'-GCTAGAATTCGGAGGCCGTCTCCGAGGA-3') and
NTR23 (5'-CGCTCTAGAAGGCGTTCCATGGCGAA-3') , respectively. NTR9, NTR16, and NTR20 contain *Eco*RI sites engineered into their 5' ends (underlined) and NTR22 and NTR23 contain *Xba*I sites engineered into their 5' ends (underlined) which allow for directional cloning into the multiple cloning site of pIJ4083. There was some evidence for a promoter within the *bldA* gene of *S. coelicolor* extending in the antisense direction

Figure III.13.1

Promoter probe plasmids for identification of the *bldA* promoter. The primers used to amplify the putative promoter fragments as well as their orientation are shown in relation to where they hybridize on the 1.5 kb *Sal*I *bldA*-containing insert in the plasmid p9S+R. Inserts 1, 2, and 3 were digested with *Eco*RI and *Xba*I and ligated into similarly digested pIJ4083. Insert 10-9 was cloned as an *Eco*RI-blunt ended fragment into *Eco*RI-*Sma*I digested pUC118. The insert was removed as an *Eco*RI-*Hin*dIII fragment and ligated into similarly digested pIJ4083. The *ermE** fragment was obtained as an *Eco*RI-*Bam*HI fragment and cloned into similarly digested pIJ4083 for use as a positive control. The vector pIJ4083 without insert was used as a negative control. The primers NTR24 and NTR26, shown in the MCS of pIJ4083, were used to amplify inserts in pIJ4083 to verify plasmid integrity.


SalI
ORF1
SmaI
NTR20
SmaI
NTR16
NcoI
NTR9
bldA
SmaI
NTR23
NTR22
NTR10
Insert 1
EcoRI
XbaI
Insert 2
EcoRI
XbaI
Insert 3
EcoRI
XbaI
Insert 10-9
EcoRI
HindIII
Insert ermE*
EcoRI
BamHI


NTR24
HindIII EcoRI SstI BamHI XbaI
NTR26
fd term
MCS
xylE
pIJ4083
7.60 Kb
neo
tsr

relative to the tRNA (Leskiw *et al.*, 1993), so this region of *S. clavuligerus bldA* was also tested for promoter activity (insert 10-9). This insert was generated by PCR using primers NTR9 and NTR10. The insert was digested with *Eco*RI and cloned into *Eco*RI-*Sma*I digested pUC118. The insert was removed as an *Eco*RI-*Hin*dIII fragment and directionally cloned into pIJ4083. The constitutively expressed *ermE** promoter was cloned into pIJ4083 as an *Eco*RI-*Bam*HI fragment, to serve as a positive control. The vector, pIJ4083, without insert was used as a negative control. The promoter probe plasmids were initially transformed into *S. lividans,* and the plasmid integrity was verified by restriction analysis or by PCR using either the primers that generated the inserts or the primers NTR24 (5'-GATCTGCAGCCAAGCTT-3') and NTR26 (5'-ATAGTTCATGTTGTCAGGT-3') which are homologous to regions flanking the MCS in pIJ4083. After isolation of plasmid DNA from *S. lividans,* the promoter probe constructs were transformed into *S. clavuligerus* NRRL 3585.

The relative promoter activity of each of the putative promoter sequences was determined by assaying C23O activity of cell free extracts of *S. lividans* and *S. clavuligerus* broth cultures. In the presence of catechol, C23O catalyzes the production of 2-hydroxymuconic semialdehyde, a bright yellow colored product which can be measured spectrometrically (described in section II.10.2).

There was considerable variation in growth rate between the various transformants, so the transformants were initially inoculated into 5 mL of TSB:+ 1% glycerol *(S. lividans)* or TSB + 1% starch *(S. clavuligerus)* broth and incubated until the cultures were sufficiently turbid. These seed cultures were then used to inoculate 25 mL of TSB + starch/glycerol broth (250 µL inoculum) and the cultures were further incubated for 24-48 hours, until there was sufficient biomass to prepare cell free extracts. Because of these variables, the promoter activity of each of the constructs cannot be correlated with stage of growth and *bldA* accumulation. However, the results are reproducible as each construct was tested twice from separate cultures (Table III.13.1).

As expected, the vector alone without insert showed low levels of activity. The construct 10-9 showed results similar to the negative control in *S. lividans,* and in *S. clavuligerus* the measurable C23O levels were even lower. This construct, which contains the tRNA gene in the reverse orientation, could be forming hairpin structures which are functioning as transcription terminators. Alternatively, the presence of insert without a promoter could be sufficient to decrease the amount of read-through transcription from the vector sequences, although the level of read-through transcription in pIJ4083 is expected to be very low due to the presence of the strong fd terminator. The *ermE** promoter showed high levels of activity in *S. clavuligerus* and somewhat lower levels in *S. lividans.* Since the *ermE** promoter is constitutively expressed, the variation is most likely due to variations in growth. *S. lividans* cultures had a tendency to grow as clumps which were quite resistant to sonication. The promoter probe constructs 1, 2, and 3 all showed C23O activity above the basal level of pIJ4083. Promoter probe construct 1 is a relatively weak promoter and based on homology to the *S. coelicolor bldA* promoter, the predicted transcription start site is 69-70 bases upstream from the tDNA. Promoter probe constructs 2 and 3, which contain overlapping fragments, also show promoter activity. In the *S. lividans* samples, construct 3 shows approximately twice the activity of construct 2, indicating that there may be two promoters in this region, one within insert 2 and one upstream of insert 2. For the *S. clavuligerus* samples,

Table III.13.1 Catechol dioxygenase (C23O) activity in cell free extracts of *Streptomyces clavuligerus* and *Streptomyces lividans* transformed with pIJ4083 recombinant plasmids for identification of the *bldA* promoter

	HOST ORGANISM					
	S. clavuligerus			*S. lividans*		
	C23O specific activity (mU/mg protein)[b]					
PLASMID[a]	Sample 1	Sample 2	Average	Sample 1	Sample 2	Average
pIJ4083	18	13	16	10	8	9
ERME*	9 851	12 245	11 048	2 002	6 246	4 124
1	74	45	60	26	35	31
2	409	449	429	148	62	105
3	563	167	350	327	386	357
10-9	2	3	3	10	11	11

[a] Recombinant plasmids as illustrated in Figure III.13.1.

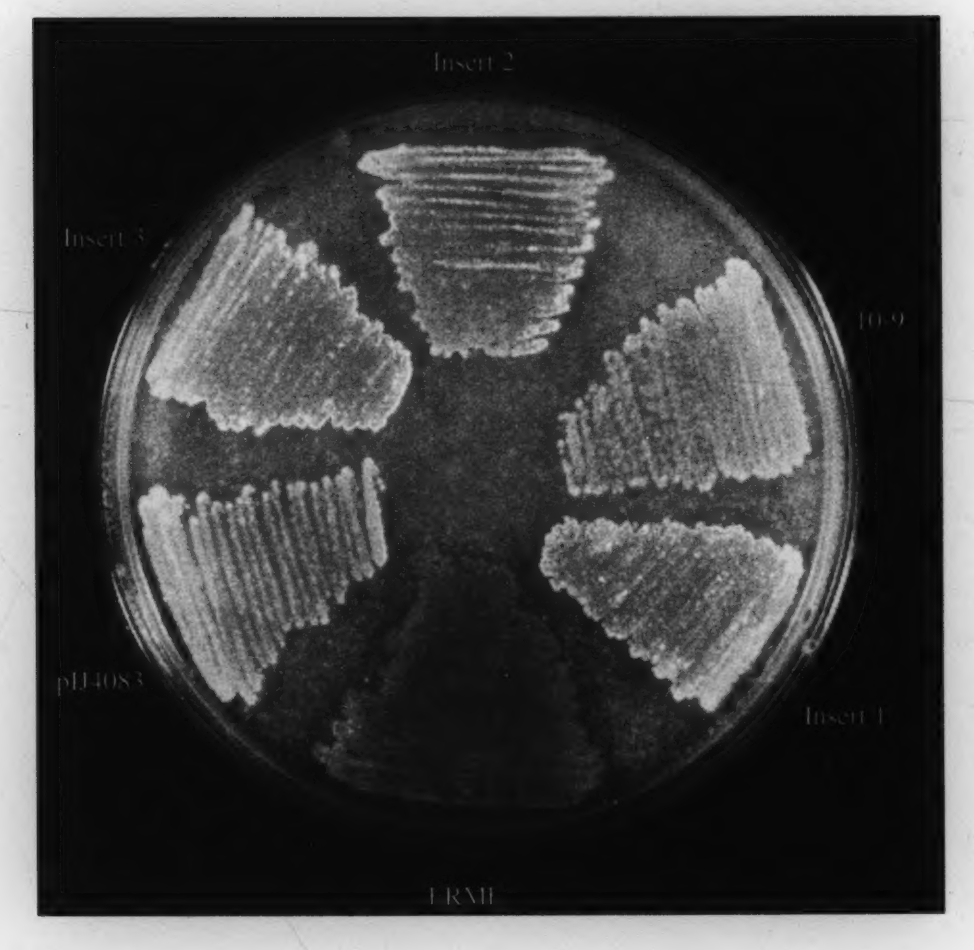
[b] C230 activity was determined by the method of Zukowski *et al* (1983) using a molar absorption coefficient of 3.3×10^4 (Sala-Trepat and Evans, 1971). One millunit (mU) corresponds to the formation at 30°C of 1 nmol of 2-hydroxymuconic semialdehyde per minute.

there is considerable variation in the activity observed for construct 3. If the promoter is temporally regulated this variation could be due to collecting the samples at different stages of growth. Another explanation is that much of the promoter activity was lost due to deletion by the time the second sample was collected as this construct was unstable. However, this second possibility was not verified by checking plasmid integrity at the time of harvest. In summary, the results demonstrate that there is a least one promoter in the region covered by construct 2 and further tests would have to be performed to determine if there is a third promoter in the region immediately upstream of construct 2. Figure III.13.2 shows *S. clavuligerus* containing the various promoter probe constructs when patched onto minimal medium + 1% maltose agar and sprayed with catechol.

Figure III.13.2

Visual assessment of catechol dioxygenase activity in patches of *S. clavuligerus* transformed with the recombinant promoter probe plasmids. The relevant insert in the vector pIJ4083 is indicated. The patch marked pIJ4083 represents the vector without insert. The cultures shown were grown on minimal medium +1% maltose agar.

S. clavuligerus NRRL 3585


Insert 2
Insert 3
10-9
pIJ4083
Insert 1
ERMII

IV. DISCUSSION

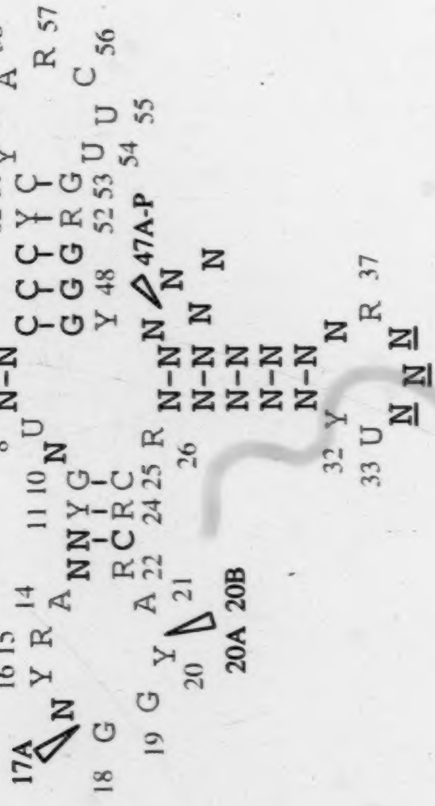
The *S. clavuligerus bldA* gene was cloned and sequenced as a 1.5 kb *Sal*I fragment. The tRNA gene is located approximately in the middle of the fragment and is flanked by divergent protein coding sequences. The two flanking, partial open reading frames show homology to genes located near the *bldA* gene of *S. coelicolor* (Lawlor, 1987) and *S. griseus* (Kwak *et al.*, 1996), indicating that this region is conserved to some extent among the streptomycetes.

One of these conserved features is a region of high G+C content containing several large, 13-20 base imperfect direct repeats located between the *bldA* gene and ORF2. The function of this region, if any, is unknown, however, it does show some similarity to a DNA fragment of *S. fradiae (micX)* that was found to activate actinorhodin production in *S. lividans* (Romero *et al.*, 1992). The *micX* sequence from *S. fradiae,* and the *S. lividans micX* homologue have an overall G+C content of 80%, with the transcribed strand containing 48% G. These activating sequences are believed to function as antisense RNA regulators, as they do not encode a protein product but are transcribed. It would be interesting to determine if this high G+C region downstream of the *bldA* gene performs a similar function. Certainly, this region has the potential to form secondary structures, and perhaps more unusual tertiary structures, which may be of biological significance. It is well-known that the species *S. clavuligerus* is highly recombinogenic (B. Barton, personal communication), perhaps this high G+C repetitive sequence is a hotspot for recombination. At any rate, the unusual nature of this sequence warrants further study.

The *bldA* gene encodes a tRNA of 83-4 nucleotides. It contains the anticodon TAA which translates leucine UUA codons in mRNA. The tDNA shows 93% homology to the *S. coelicolor* (Lawlor *et al.*, 1987) and *S. lividans bldA* (Ueda *et al.*, 1992) tDNAs, and 92% homology to the *S. griseus bldA* tDNA (Kwak *et al.*, 1996). The sequence of 50 *Streptomyces* tDNAs were compared to identify the invariant and semi-invariant bases in *Streptomyces* tRNAs (Table III.4.1 and Table III.4.2). From this comparison the *Streptomyces* invariant and semi-invariant bases were identified and are shown in Figure IV.1. The absence of the G18 nucleotide in the D-loop of *bldA* tRNAs is highly unusual as this nucleotide is highly conserved in all cytoplasmic tRNAs from all organisms, including *Streptomyces* tRNAs. The GG (G18-G19) sequence in the D-loop has been proposed to stabilize the three-dimensional structure of the tRNA by forming hydrogen bonds with pseudouridine (U55) and cytidine (C56) in the T-loop, so the *bldA* tRNAs may form altered, less stable three-dimensional structures. Consistent with this prediction, the unique *bldA* D-loop structure was found to affect the structural stability of the tRNA, as it was found to be highly susceptible to temperature changes in the absence of Mg^{2+} (Ueda *et al.*, 1992). The altered D-loop structure was also found to have altered sensitivity to nuclease *S1* and RNase T1 when limited digestion of the *S. lividans bldA* tRNA was compared to an engineered *bldA* tRNA containing an *E. coli* $tRNA^{Leu}_{UAA}$ D-loop (9 bases including G18-G19), indicating that the unique D-loop induces conformational changes on the tRNA (Ueda *et al.*, 1992). While the tRNAs containing the unique D-loop structure had a lower affinity for the *E. coli* leucyl-tRNA synthetase (LeuRS), this did not appear to be the case with the *S. lividans* LeuRS, suggesting that *Streptomyces* LeuRS has a unique recognition mechanism toward *bldA* tRNAs. If this is the case then

Figure IV.1

Cloverleaf structure of *Streptomyces* tRNAs, showing the semi-invariant and invariant bases. An N indicates any nucleotide, an R indicates a purine (G or A), and a Y indicates a pyrimidine (U or C). The location of additional bases that may be present (17A, 20A, 20B in the D-loop, and 47A-P in the variable loop) are indicated. Semi-invariant and invariant bases are indicated in grey, the anticodon is underlined.


3'
A
C
C 74
5'
N-N
N-N
N-N
N-N
N-N
N-N
N-N
N
U
8
N
11 10
N N Y G
C R C
R
24 25
R
26
A 14
16 15
Y R
17A
N
18
G
19
G
Y
20
21
20A 20B
A 22
C C C
G G G
Y 48
47A-P
N
62 61
60
Y
N
A 58
R 57
C 56
U 55
U 54
G
R
52 53
N-N
N-N
N-N
N-N
N-N
N
N
N
32 Y
33 U
N N N
R 37

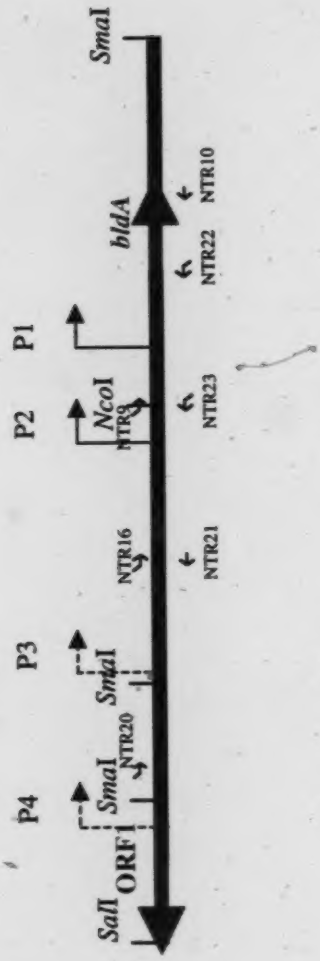
Streptomyces tRNA

this may be a mechanism by which the unique D-loop structure functions to increase the amount of charged *bldA* tRNA compared to other leucyl-tRNA species late in growth, which may have implications in the translation of *bldA* targets. However, a better understanding of the role of the unique D-loop of *bldA* tRNAs requires additional studies that compare the aminoacylating activity of *Streptomyces* LeuRS with other *Streptomyces* leucyl-tRNA species, and the interactions of leucyl-tRNAs with EF-Tu to determine the role of *bldA* tRNA in protein biosynthesis.

Overall, the results from *S1* nuclease protection analysis, primer extension analysis and promoter probe analysis suggest that the *bldA* gene of *S. clavuligerus* contains at least two promoters and may contain as many as four promoters (Figure IV.2). It is surprising that the *S1* nuclease protection studies showed full-length protected fragments without identifying any primary transcripts. Although it is possible that these full-length protected fragments are artifacts caused by 'breathing' of probe-probe reannealed fragments at the 5' end, this seems unlikely since promoter probe studies also indicate that there are additional promoters, located upstream of the putative promoter that shows homology to the *S. coelicolor bldA* promoter. Also, the absence of these full-length protected fragments in the *bldA* strain suggests that they are not artifacts. Alternatively, while it is possible that the primary transcripts were processed too quickly to be detected by *S1* nuclease mapping, this also seems unlikely since there was sufficient primary transcript to detect full-length protected fragments. At any rate, additional tests are required to determine the exact number and location of the *S. clavuligerus bldA* promoters. Some of these experiments should include: *S1* nuclease protection studies on *bldA* transcripts from a strain carrying the *bldA* gene in high copy number; additional

Figure IV.2

Schematic diagram of the *bldA* region with the putative promoter consensus sequences indicated. The location of promoter P1 is 3' of the *Nco*I site, and based on its homology to the *S. coelicolor bldA* promoter and the location of promoter consensus sequences (Bourn and Babb, 1995) it is most likely located at the position indicated. The location of promoter P2 was identified by promoter probe studies to lie between primers NTR16 and NTR23 and based on the location of promot[illegible]onsensus sequences it most likely is located where indicated. There is some evidence for promoters P3 and P4 and their approximate positions are indicated.


P4
P3
P2
P1
SalI
ORF1
SmaI
NTR20
SmaI
NTR16
NcoI
NTR8
bldA
SmaI
NTR21
NTR23
NTR22
NTR10

promoter probe studies with fragments in the -334 to -506 (putative promoter 3) region and -506 to -681 (putatative promoter 4) region to determine if there is a third and fourth promoter. These new constructs, as well as the previously tested constructs, should be assayed for activity throughout growth to determine when each promoter is active as well as to determine the relative strength of each promoter. Although promoter probe analysis indicated that there is no antisense promoter located within the *S. clavuligerus bldA* tRNA, it would be worthwhile to determine if there is an antisense promoter located 3' of the *bldA* tRNA as sequence analysis indicates that there are several promoter consensus sequences (Bourn and Babb, 1995) in this region.

The regulation of tRNA gene transcription by more than one promoter is not unusual. For example, the *S. lividans glyUβ* tRNA gene is transcribed by dual promoters (Sedlmeier *et al.*, 1993) and the $\text{tRNA}^{\text{Val}}_{\text{GAC}}$ in the carrot chloroplast rDNA cluster is transcribed by three promoters (Manna *et al.*, 1994). The $\text{tRNA}^{\text{Val}}_{\text{GAC}}$ gene, is temporally regulated, similar to the pattern of accumulation observed for the mature, processed form of the *bldA* tRNA. tRNA transcripts are not detected during early embryonic stages and progressively accumulate during late embryonic stages. It is also not uncommon for tRNA genes associated with rRNA operons to be transcribed by their own promoter as well as cotranscribed with the rRNA (Vold *et al.*, 1988a). It has been suggested that the clusters of tRNA genes in *E. coli* contain only 5-7 tRNA genes (as opposed to *Bacillus subtilis* tRNA clusters which contain up to 27 tRNA genes) to allow for heterogeneous regulation of the clusters (Rowley *et al.*, 1993). The clusters usually consist of tRNA species that perform similar functions, for example, translate major codons, and the clustering of the tRNA genes may allow for co-ordinate expression. When the promoter regions of four *E. coli* tRNA operons containing leucyl-tRNAs were analyzed it was found that the four promoter regions had different responses to growth rate, stringent conditions induced by serine hydroxamate, analog inhibition, or leucine starvation (Rowley *et al.*, 1993). A similar situation may exist in *Streptomyces,* as tRNA genes can occur singly, or in clusters of up to five tRNA genes. The tRNA genes that are co-transcribed can be co-ordinately regulated while the tRNA genes that are transcribed as mono-cistronic transcripts, such as the *bldA* tRNA genes may be differentially regulated. tRNA genes encoding tRNA species that translate minor codons, for instance, generally are not as highly responsive to stringent conditions as are the genes that encode tRNA species that translate major codons. For example, the promoter for the *leuX* cluster which translates the minor $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{CAA}}$ of *E. coli* decreased transcription only 3-fold in response to stingent conditions whereas promoter acitivity of operons containing major leucyl-tRNAs decreased transcription 10-fold (Rowley *et al.*, 1993). The likely reason for this differential response to stringent response induction is that genes expressed during exponential growth consist primarily of major codons and the genes containing minor codons are usually confined to genes expressed during periods of slower growth, or periods when the stringent response is expected to reduce the transcription of major tRNA species. The strong repression of major tRNA gene transcription decreases the total tRNA pools while the relative concentration of the minor tRNA species needed to translate the minor codons increases. Consistent with this differential regulation, most of the tRNA genes that translate rare codons are transcribed alone. This includes *bldA* of *Streptomyces, leuX* of *E. coli,* which is implicated in fimbrial phase variation (Newman *et al.*, 1994), $\mathrm{tRNA}^{\mathrm{Thr}}_{\mathrm{CGU}}$ of *Clostridium acetobutylicum* which is implicated in solventogenesis (Sauer and Dürre, 1992), and *dnaY* ($tRNA^{Arg}_{UCU}$) of *E. coli* (Garcia *et al.*, 1986). As these tRNAs are all required at a specific point in development or growth, it makes sense that they are transcribed independently. So although it is not unusual for a tRNA to be regulated by a number of promoters it is unusual that the *S. coelicolor* and *S. clavuligerus bldA* tRNAs are apparently regulated by a different number of promoters. This is unexpected since the accumulation of the *bldA* tRNAs in both organisms is similar, as well as the distribution of TTA codons, indicating that the *bldA* tRNAs perform a similar role and would be expected to be subject to similar regulatory mechanisms. Perhaps the additional promoters regulating the *S. clavuligerus bldA* tRNA are involved in sensing and responding to environmental conditions, such as stress. Thus far, experiments have not addressed what effect stress and other environmental conditions have on *bldA* expression. The experiments in *S. coelicolor,* by Gramajo *et al.*, (1993), indicate that at least under some conditions, functional*bldA* tRNA is present early in growth, suggesting that *bldA* gene expression in *S. coelicolor* may be responsive to different environmental conditions. Perhaps *S. clavuligerus* is similarly responsive to environmental conditions, but differs from *S. coelicolor* in the mechanism of regulation of *bldA* gene expression. At any rate, much remains to be learned about the regulation of expression of *bldA* in both *S. coelicolor* and *S. clavuligerus.*

An *S. clavuligerus bldA* mutant strain was generated by gene replacement, and as expected the mutant colonies were sporulation defective. Surprisingly, the mutant colonies were not deficient in antibiotic production as both cephamycin C and clavulanic acid were produced. This result was unexpected since previously characterized *bldA* mutants from *S. coelicolor* and *S. griseus* were defective in antibiotic production, and the transcriptional activator, (CcaR), for the cephamycin C and clavulanic acid biosynthetic pathways, in *S. clavuligerus* (Perez-Llarena *et al.*, 1997), is a *bldA* target. This raises the question as to how cephamycin C and clavulanic acid are produced in the *bldA* mutant strain when synthesis of these compounds should be dependent on the *bldA* tRNA. The most obvious possibility is that there is more than one gene that encodes $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{UAA}}$, however, there are several lines of evidence that indicate that this is not the case. First of all, the *S. coelicolor bldA* gene hybridized to a single gene in the *S. clavuligerus* genome, under conditions of low stringency. If there was a second *bldA* gene in the *S. clavuligerus* genome it should hybridize to the *S. coelicolor bldA* gene under these conditions. Secondly, there was no detectable *bldA* tRNA in RNA isolated from the *bldA* mutant strain. The probe used to detect the *bldA* tRNA was a 17 mer synthetic oligonucleotide BKL5 (5'-TTAAGCTCGCCGTGTCT-3') which is complementary to the anticodon stem loop and to the D stem. Hybridizations were often carried out at 37°C, which is 15° below the T_d of the primer and would detect tRNAs with two mismatches. While it is possible that an additional $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{UAA}}$ could contain more than two base differences over this region it seems unlikely since all of the *bldA* genes sequenced to date are conserved over this region. However, the most compelling evidence that there is a single copy of $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{UAA}}$ in the *S. clavuligerus* genome is that deletion of the *bldA* gene resulted in a sporulation defective phenotype. If there was a second *bldA* gene then deletion of one of the copies should not result in a noticeable phenotype. Moreover, the *bldA* gene in single copy complemented the mutant phenotype as it restored sporulation. In addition, the reporter gene *gfp*, which contains three TTA codons, was expressed in the wild type *S. clavuligerus* strain but not the *bldA* mutant strain, which indicates that the *bldA* mutant strain cannot efficiently translate the TTA codons in this reporter gene. All of these lines of evidence taken together suggest that there is only one gene encoding $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{UAA}}$.

Another possible reason that cephamycin C and clavulanic acid are produced in a *bldA* mutant could be that a transcription factor other than CcaR activates the biosynthetic pathways. The biosynthesis of clavulanic acid and clavam compounds is similar, and several of the biosynthetic genes of these two pathways are homologous, so it is possible that the activator for the clavam biosynthetic pathway might activate the clavulanic acid pathway (and subsequently activate the cephamycin C pathway), under some conditions. Again, there are several lines of evidence that indicate that this is not the case. A *ccaR* mutant strain is defective in both clavulanic acid and cephamycin C production (Perez-Llarena *et al.*, 1997) (Alexander and Jensen, 1998), indicating that the CcaR protein is required for the activation of these two biosynthetic pathways. The *bldA* mutant strain produces cephamycin C and clavulanic acid when grown on starch asparagine medium, conditions under which clavam compounds are not produced (Paradkar and Jensen, 1995), and western blot analysis indicates that CcaR protein is produced in the *bldA* mutant strain under these conditions. All of these results indicate that CcaR is the only activator of the cephamycin C and clavulanic acid biosynthetic pathways, and is produced in the *bldA* mutant strain.

If there is only one copy of the *bldA* gene and the *bldA* target CcaR is required for cephamycin C and clavulanic acid production, then translation of the single UUA codon in the *ccaR* transcript must be by a non-cognate tRNA in the mutant strain. While there is no direct evidence to support this hypothesis there are many precedents in the literature that indicate that mistranslation (translation of a codon by a non-cognate tRNA) can and does occur frequently. Historically, the earliest and perhaps best-known examples of mistranslation are by suppressor tRNAs which translate nonsense and missense mutations as well as frameshift mutations (Murgola, 1995). Suppressor tRNAs are created by mutations to a tRNA such as $\text{tRNA}^{\text{Leu}}_{\text{CAA}}$ (Thorbjarnardottir *et al.*, 1985) which enables them to translate missense, frameshift or nonsense mutations, such as those encoding UAG amber stop codons. In most of these cases, mistranslation is the result of selective pressure, and the suppressor tRNA results from a secondary mutation that is generated in response to the primary mutation that initially created the nonsense, missense, or frameshift mutation. Although suppressor tRNAs can mistranslate efficiently, they are not the most relevant example of mistranslation as they are usually created in response to strong selective pressure, and are typically not present under normal (wild-type) conditions. What is more relevant is the occurrence of low-level suppression of missense, nonsense, and frameshift mutations in wild-type systems (Farabaugh and Bjork, 1999; Kurland and Ehrenberg, 1984). Since suppressor tRNAs are presumably not present in wild-type systems, suppression of these missense, nonsense, and frameshift mutations under these conditions indicates that wild-type tRNAs are capable of translating non-cognate codons. It is estimated that the frequency of translational errors is 10^{-5} to 10^{-4}, although under some conditions it can be as high as 10^{-2} (Kurland and Ehrenberg, 1984). In some cases frameshifting or read-through of nonsense codons acts as a regulatory device, and mistranslation at these 'programmed sites' can approach 100% efficiency (Farabaugh and Bjork, 1999). For example, the *gag-pol* shift in HIV-1 (Jacks *et al.*, 1988b) and the Rous Sarcoma virus (Jacks *et al.*, 1988a) is caused by slippage of the translational apparatus to the –1 frame at specific sequences, and expression of the yeast Ty1 and Ty2 elements requires a similar shift to the +1 frame at specific sites (Belcourt and Farabaugh, 1990). However, in many cases the low-levels of suppression in wild-type systems are indicative of very high levels of mistranslation. In these cases optimization of the growth rate occurs at the expense of translational accuracy.

The mechanics of mistranslation are complex, there are several factors other than growth rate (intracellular conditions) that affect the frequency, such as tRNA modification, wobble base pairing, and codon context. Part of the maturation process of a primary transcript to a mature, charged-tRNA includes modification of specific bases. Several of the modified bases that have been identified (D, dihydrouridine; Ψ, pseudouridine; Um, 2'-*O*-methyluridine; ac^{4}C, N^4-acetylcytidine; Cm, 2'-*O*-methylcytidine; m^{1}G, 1-methylguanosine; m^{7}G, 7-methylguanosine; Gm, 2'-*O*-methylguanosine; m^{1}A, 1-methyladenosine; t^{6}A, *N*^{6}-threonlycarbamoyladenosine; mt^{6}A, methyl-*N*^{6}-threonylcarbamoyladenosine; and I, inosine) are present in all three phylogenetic domains (Archaea, Bacteria, and Eucarya), and some are even present in comparable positions in the tRNAs from the three phylogenetic domains, indicating that these modified bases were likely present in the tRNA of the progenitor (Bjork, 1995). In all, more than 80 modified nucleotides have been identified in tRNA to date (see Appendix 2 for a list of modified nucleotides in tRNAs). The function of the modified nucleotides varies; some may function to varying degrees as identity determinants and antideterminants for aminoacylating enzymes (Bjork, 1995); some modifications may play a role in sensing and responding to environmental conditions (Kitchingman and Fournier, 1975); some modifications affect codon recognition and translation rates

(Persson, 1993), while some modifications may have dual functions (Bjork, 1995). Most of the modifications that affect codon recognition are confined to the anticodon loop. While it is known that modified nucleotides outside of the anticodon loop can alter the conformation of the tRNA molecule and affect the codon-anticodon interaction, there is insufficient data regarding the specifics of these alterations to propose how this may relate to mistranslation of a specific codon, and will not be discussed further.

Within the anticodon loop, modified nucleotides have been found to significantly affect codon recognition. The first base of the anticodon (position 34) is most often modified, to either expand the decoding capacity of the tRNA molecule, or to limit it. Because there are 64 different codons and only 20 amino acids, most of the amino acids are specified by more than one codon. The amino acids glycine, alanine, threonine, proline, and valine are specified by four codons. The four codons for each of these amino acids differ from each other by only one base, the third base in the codon. The result is that GGN encodes glycine (where N is C,G, A and U), GUN encodes valine, CCN encodes proline and ACN encodes threonine. The pattern of codon assignment is similar for the other amino acids. Amino acids encoded by six codons, such as leucine, have the 'family box' assignment (CUN) as well as the two additional codons (UUA and UUG). Amino acids encoded by two codons are specified for by either NNY (pyrimidine) or NNR (purine). The only codon assignments that differ from this pattern are the codons for methionine, tryptophan and isoleucine. Isoleucine is encoded by three codons (AUU/C/A) and the forth codon in this set, AUG is assigned to methionine. Tryptophan is encoded by UGG and its counterpart UGA encodes a stop codon (see Appendix 1 for codon designations). Thus, the degeneracy of the genetic code is mainly confined to the third codon position. It is apparent that translation rates would be optimized if the base pairing between the third base of the codon and first base of the anticodon were relaxed to enable a tRNA molecule to recognize all of the codons in a family box or the NNA/G or NNU/C pairs. This observation led Crick to purpose the Wobble Hypothesis in 1966. The basis for this hypothesis is that the first base of the anticodon may pair with more than one type of base. In order for this non-conventional pairing to occur and for two or more hydrogen bonds to form, the base in the first position of the anticodon must change its position relative to the third base of the codon; the first base of the anticodon in effect must wobble. According to this hypothesis a G in the anticodon can pair with U as well as C; U can pair with A and G; and I (inosine) can pair with U, C, and A. While Crick's hypothesis was proved to be essentially correct it didn't account for the ability of A to pair with U, C, G and A; and for U to pair with U and C as well as A and G. Crick's hypothesis also didn't account for the role that modified bases, other than inosine, might play in wobble basepairing. While some modified bases such as xo^5U (5-hydroxyuridine) at position 34 enhances the efficiency of recognition of U, A, and G; others such as mnm^5s^2U (5-methylaminomethyl-2-thiouridine) function to restrict pairing to A and G (Yokoyama and Nishimura, 1995). Modifications that enhance non-conventional pairing such as xo^5U (5-hydroxyuridine) are typically found in tRNAs specific to amino acids that are encoded with a set of four codons such as valine, serine, and alanine, whereas modifications that restrict pairing such as mnm^5s^2U (5-methylaminomethyl-2-thiouridine) are typically found in tRNAs specific to amino acids that are encoded with a set of two codons such as glutamine, lysine and glutamate.

The wobble base pairing proposed by Crick as well as the recently revised wobble base pairing is shown in Table IV.1. According to the revised wobble rules, the *bldA* target UUA could theoretically be translated by the anticodons AAA, k^2CAA, or IAA, in addition to its cognate anticodon UAA. However, not all of these possibilities are likely. It is unlikely that tRNAs with either AAA or IAA anticodons translate the UUA codon. An unmodified wobble A in the P-site tRNA weakens the codon-dependent binding of the incoming A-site tRNA by interduplex interaction and leads to a low efficiency of ribosomal translation (Lim, 1995). As such, unmodified A is rarely present in position 34 of the anticodon and has only been found in a $\text{tRNA}^{\text{Arg}}_{\text{ACG}}$ from yeast mitochondria (Bjork, 1995) and $\text{tRNA}^{\text{Thr}}_{\text{AGU}}$ from *Mycoplasma* spp. (Inagaki *et al.*, 1995), so it is unlikely that the A34 of the *S. clavuligerus* $\text{tRNA}^{\text{Phe}}_{\text{AAA}}$ is unmodified. Formation of inosine often occurs through the deamination of A, however, this modification has only been found in tRNAs that decode family boxes, so again it is unlikely to be the modified base present in $\text{tRNA}^{\text{Phe}}_{\text{AAA}}$. It is impossible to predict what modification occurs to the A34 of $\text{tRNA}^{\text{Phe}}_{\text{AAA}}$ in *Streptomyces*, however, the modification likely limits its decoding capacity to UUU and possibly UUC phenylalanine codons (Curran, 1995; Wilson and Roe, 1989). If the other phenylalanyl-tRNA ($\text{tRNA}^{\text{Phe}}_{\text{GAA}}$) contains an unmodified G34, it will also be limited to decoding UUU and UUC codons. Most eukaryotic and prokaryotic $\text{tRNA}^{\text{Phe}}_{\text{GAA}}$ species contain a ribose methylated nucleoside at position 34 (Gm34) however, this modification may contribute to prohibition of noncognate (UUA and UUG) codon reading, so it is unlikely that this phenylalanyl-tRNA can mistranslate a UUA codon whether the G34 is modified or not. The only likely possibility for mistranslation of UUA codons by wobble base pairing at position 34 is by $\text{tRNA}^{\text{Leu}}_{\text{CAA}}$. The 34C of most UUG-reading leucyl-tRNA

Table IV.1 Wobble base pairing between the first nucleotide of the anticodon and the third nucleotide of the codon

Crick's wobble rule		Revised wobble rule	
First nucleotide of anticodon	Third letter of codon	First nucleotide of anticodon	Third letter of codon
G	U, C	G	U, C
C	G	C	G
		k^2C	A
		A	U, C, G, (A)
U	A, G	U	U, A, G, (C)
		xm^5s^2U, xm^5Um, Um, xm5U	A, (G)
		xo^5U	U, A, G
I	U, C, A	I	U, C, A

k^2C denotes lysidine. xm^5s^2U, xm^5Um, and xm^5U denote 5-methyl-2-thiouridine, 5-methyl-2'-*O* -methyluridine, and 5-methyluridine derivatives, respectively, where "xm" denotes, e.g., methylaminomethyl (mnm), carboxymethylaminomethyl (cmnm), or methoxycarbonylmethyl (mcm). xo^5U denotes a 5-hydroxyuridine derivative such as 5-methoxyuridine (mo^5U) and 5-carboxymethoxyuridine (cmo^5U).

Copied from Yokoyama and Nishimura (1995) species has been found to contain 2'-O-methylcytidine (Cm) (Yokoyama and Nishimura, 1995), and in *E. coli* the CmAA anticodon is able to recognize and translate the UUA codon under certain conditions (Takai *et al.*, 1994). In the study conducted by Takai et al. (1994), the translation of three UUA codons in the chloramphenicol acetyltransferase (CAT) protein was measured by incorporation of $[^3H]$leucine. In competitive reactions using $[^3H]$leucine-charged $\text{tRNA}^{\text{Leu}}_{\text{CAA}}$ and unlabelled $\text{tRNA}^{\text{Leu}}_{\text{UAA}}$, there was little incorporation of radioactively labeled leucine into the CAT protein. However, in the absence of the competitive $\text{tRNA}^{\text{Leu}}_{\text{UAA}}$ species, there was significant incorporation of $[^3H]$leucine indicating that under non-competitive conditions, the $\text{tRNA}^{\text{Leu}}_{\text{CAA}}$ can efficiently translate UUA codons in *E. coli.* Although it is currently not known whether the 34C of *Streptomyces* $\text{tRNA}^{\text{Leu}}_{\text{CAA}}$ is similarly modified, this tRNA species seems a likely candidate for mistranslation of the UUA codon in the *ccaR* transcript in the *S. clavuligerus bldA* mutant strain.

Position 37 of the anticodon loop is also highly modified in tRNAs (Bjork, 1995). Although this nucleotide lies outside of the anticodon (34-36), it does play a role in codon-anticodon recognition. It is believed that modification of this nucleotide has evolved to strengthen the base pairing between position 36 (third base of the anticodon) and the first base of the codon (Yokoyama and Nishimura, 1995). In keeping with this theory the type of modification to the nucleotide in position 37 correlates with the kind of nucleotide present in position 36. tRNAs that recognize codons starting with U and consequently have an A in position 36 usually contain a bulky hydrophobic modified nucleotide in position 37, such as yW (wybutosine) or i^6A (isopentenyladenosine) derivatives. These modifications to nucleotide 37 may improve the A36-U interaction and also prevent A36 from base pairing with bases other than U (Bjork, 1995). Although the base modifications present in the *Streptomyces bldA* tRNAs have not yet been determined, all of them contain A37, likely candidates for i^{6}A modifications.

tRNAs that decode codons starting with C have m^{1}G (methylguanosine), m^{2}A (methyladenosine), or in a few cases unmodified A in position 37 (Yokoyama and Nishimura, 1995). The m^{1}G37 modification is highly conserved; it is present in tRNAs (the subset of tRNAs that read codons starting with C) from all three phylogenic domains which suggests that this modification was also present in the progenitor. All tRNAs that read codons CUN (leucine), CCN (proline), or CGN (arginine) contain m^{1}G37, and consistent with this, the two tRNA $^{\text{Leu}}_{\text{GAG}}$ identified in *S. coelicolor* contain G37. The m^{1}G37 nucleotide is believed to prevent frameshifting and may also improve the cognate interaction. tRNAs that read codons starting with G frequently have unmodified nucleotides at position 37. If nucleotide 37 is modified it is often modified to methylated adenosine derivatives (m^{2}A and m^{6}A), methylated guanosine (m^{1}G), or methylated inosine (m^{1}I). Codons that start with A are decoded by tRNAs that typically have a t^{6}A (threonylcarbamoyladenosine) derivative in position 37, the most notable exceptions are initiator tRNA$^{\text{fMet}}$ which contain an unmodified A37. It is believed that the lack of modification to A37 may allow U36 to wobble, which allows tRNA$^{\text{fMet}}$ to decode AUG, GUG, UUG, and CUG codons. An alternative theory for the ability of the initiator tRNA to recognize NUG codons is that it is a function of the ribosomal P site. However, there are additional examples of wobble base pairing with nucleotide 36 that occur with tRNAs other than fMet and thus would occur in the ribosomal A site making it unlikely that the wobble base pairing of fMet is solely a function of the ribosomal P site.

An example of wobble base pairing of nucleotide 36 is the frameshifting of the Ty1 and Ty2 elements in yeast (Belcourt and Farabaugh, 1990). This frameshifting involves an unusual $\text{tRNA}^{\text{Leu}}_{\text{UAG}}$ that is able to recognize all six leucine codons (CUN and UUA/G) in the sequence CUU AGG C. The $\text{tRNA}^{\text{Leu}}_{\text{UAG}}$ initially pairs with the CUU codon, then slips to pair with the UUA leucine codon in the +1 frame. The ability of $\text{tRNA}^{\text{Leu}}_{\text{UAG}}$ to pair with both CUU and UUA leucine codons indicates that G36 of the anticodon may be allowed to wobble to pair with U as well as C, and U34 can wobble to pair with U as well as A. If wobble base pairing can occur with nucleotide 36 of the anticodon and the first base of the codon then UUA codons could theoretically be translated by UAN containing anticodons $\text{tRNA}^{\text{Leu}}_{\text{UAG}}$, $\text{tRNA}^{\text{Ile}}_{\text{UAU}}$, and $\text{tRNA}^{\text{Val}}_{\text{UAC}}$. It seems improbable that a significant amount of wobble base pairing with nucleotide 36 would be allowed, as this would cause the incorporation of the incorrect amino acid in most instances. However, in the absence of a cognate tRNA, wobble base pairing between position 36 and the first codon position could allow translation to continue. The incorporation of a single incorrect amino acid into a protein seems favorable over the alternative, which is the formation of a truncated protein. Although there is little evidence to indicate that this occurs, it cannot be discounted as a possible means by which the UUA codon in the *ccaR* transcript is translated in the *S. clavuligerus bldA* mutant. The possibility also exists that wobble base pairing between the first codon position and third anticodon position is limited to tRNAs that are specified by six codons, such as $\text{tRNA}^{\text{Leu}}_{\text{UAG}}$ and $\text{tRNA}^{\text{Arg}}_{\text{UCG}}$, where wobble base pairing between the first codon - third anticodon, and third codon - first anticodon positions would allow a single tRNA to read all six cognate codons. While it is evident that most modifications to nucleotide 37 prevent codon recognition by non-cognate tRNAs, very little is know about what modifications are present in *Streptomyces* tRNAs, and what effect they have on the decoding capacity of *Streptomyces* mRNA.

In contrast to the large number of modifications that occur to position 34 and 37 of the anticodon loop, there are very few modifications that occur to position 35 and there is no evidence that position 36 is modified at all. The only modification that occurs to position 35 is Ψ (pseudouridine). This modification has been found in $tRNA^{Tyr}$ and in amber (CΨA)- and ochre (UΨA)-suppressor tRNAs. These modifications are believed to stabilize base pairing with A in the second codon position (Bjork, 1995). There is no evidence to suggest that base 35 is capable of wobble base pairing, and logically wobble base pairing or mispairing at position 35 would be detrimental as it would allow the incorporation of amino acid by a non-cognate tRNA. It is therefore unlikely that the UUA codon in the *ccaR* transcript can be translated by a $tRNA^{Ser}_{UGA}$ that would contain a mismatch at position 35. The other two codons that differ from the UUA codon at the middle position specify ochre (UAA) and opal (UGA) stop codons.

The relative position of a particular codon in the mRNA, as well as the distribution of codons and nucleotides in the vicinity of the codon of interest can significantly affect translation and mistranslation rates. The best example of how codon context affects mistranslation involves frameshifting. In some cases, the expression of a gene requires a frameshift for expression, for example, many retroviral genes such as *gag-pol* in HIV-1 require a –1 frameshift for expression (Atkins and Gesteland, 1995; Jacks *et al.*, 1988b). However, in an ever increasing number of cases it appears that frameshifting occurs at specific sequences and yields non-functional products. Whether these cases are ways a cell can cope with tRNA imbalance or starvation for a single amino acid is debatable, however, the frequency of frameshifting at particular sequences is responsive to intracellular conditions. There are a surprising variety of mechanisms that result in frameshifts. In some instances, the shift is caused by slippage of tandem tRNAs (Atkins and Gesteland, 1995). The codons that cause this shift are typically X XXY YYZ, where the first and second nucleotides of the two codons are the same in the zero and –1 frames (Brierley *et al.*, 1992; Jacks *et al.*, 1988a; Jacks *et al.*, 1988b). It has been proposed that the two adjacent tRNAs pairing with the zero frame codons simultaneously slip by one nucleotide to the –1 frame, however, there are alternative theories that propose that the shift occurs after transpeptidation and during translocation. Frameshifts to the –1 and –2 positions can also occur due to slippage of a single tRNA. The slippage of the single tRNA usually occurs at runs of repeated bases such as CGG GGG, particularly when these homopolymeric runs are bordered on their 3' side by a stop codon. Although this type of frameshifting has been primarily characterized in *lacZ* plasmid constructs in *E. coli*, there is some evidence that it occurs in wildtype systems (Atkins and Gesteland, 1995). Frameshifts to the –1 position can also be created by a single tRNA without slippage. Some tRNAs, such as the $\text{tRNA}^{\text{Ser}}_{\text{GCU}}$ and $\text{tRNA}^{\text{Thr}}_{\text{GGU}}$ of *E. coli*, are inherently "shifty". It is believed that the anticodons of these "shifty" tRNAs form two-base pairs with GCA (alanine) and CCG (proline) codons respectively, to induce the –1 frameshift (Atkins *et al.*, 1979; Dayhuff *et al.*, 1986; Weiss *et al.*, 1988). There are also several types of frameshifting that result in +1 reading frame shifts. The +1 frameshift that occurs with the yeast transposable elements Ty1 and Ty2 involves repairing (or slippage) of the tRNA from the zero frame codon to the +1 frame codon

(Belcourt and Farabaugh, 1990). The sequence responsible for this frameshift is CUU AGG C. The AGG arginine codon in yeast is rare, so the corresponding $\mathrm{tRNA}^{\mathrm{Arg}}_{\mathrm{CCU}}$ is present in low amounts. The prevailing theory is that the low availability of the $\mathrm{tRNA}^{\mathrm{Arg}}_{\mathrm{CCU}}$ causes the ribosome to pause when it reaches the AGG codon. The unusual $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{UAG}}$ which is able to pair with all six leucine codons initially is paired with the CUU codon, when the ribosome pauses it shifts to pair with the UUA leucine codon in the +1 frame. The next codon becomes a GGC glycine codon and the +1 frameshift is complete. This example illustrates two important points. First of all, it suggests that a $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{UAG}}$ can translate a UUA codon under certain conditions, which may have implications for the mistranslation of the UUA codon in the CcaR protein of *S. clavuligerus*. Secondly, it suggests that the presence of a rare codon with correspondingly rare cognate tRNA (a "hungry" codon) can induce mistranslation. Frameshifts to the +1 position can also occur without repairing, particularly when the codons in the zero and +1 frames differ significantly (Farabaugh *et al.*, 1993). In *E. coli*, under conditions of lysine starvation, +1 frameshifts occur at the sequence AAG C at a high frequency (Lindsley and Gallant, 1993). Under these conditions there is a limited amount of charged lysyl-tRNA that can translate AAG codons. When the +1 frame encodes a AGC (serine) codon the "shifty" $\mathrm{tRNA}^{\mathrm{Ser}}_{\mathrm{GCU}}$ generates a +1 frameshift. This shift also occurs at lower levels at AAG U sequences but not AAG A or AAG G sequences, indicating that sequences 3' of the codon are important for shifting to occur. There are instances where the sequence to the 5' of the "hungry" codon is important as well as instances where sequences both 5' and 3' of the "hungry" codon are important for frameshifting (Lindsley and Gallant, 1993). Frameshifting also occurs at rare codons. This differs from amino acid starvation as the tRNA itself is limiting and the stringent response is presumably not involved. An example of a rare codon inducing frameshifts is the AGG (arginine) codon of *E. coli* (Rosenberg *et al.*, 1993). The presence of tandem AGG codons in the sequence AGG-AGG-U results in 50% ribosomal frameshifting (Spanjaard and van Duin, 1988). This frameshifting occurs exclusively in the +1 frame, and is suppressed by the presence of additional copies of the *argU* tRNA, indicating that the depletion of the *argU* tRNA causes the frameshift. This example is most like the conditions expected in a *bldA* mutant, where the $\text{tRNA}^{\text{Leu}}_{\text{UAA}}$ is absent. While it is apparent that the presence of the "hungry" UUA codon in the *ccaR* transcript in *S. clavuligerus* does not induce a frameshift as a functional CcaR protein is produced, it does raise the possibility that some of the UUA codons in *bldA*-dependent targets may cause frameshifts to occur. Although frameshift products have not been detected in *bldA* mutants, all of the studies to date test for target activity or the detection of target protein by western analysis. A frameshift product, if produced, is not likely to be either functional or detectable by Western analysis as the amino acid sequence, protein structure, and size would be significantly altered.

The relative position of a codon within the mRNA also can have an effect on the translation and mistranslation rates. It has been well documented that minor or rare codons are translated at a slower rate than are major codons (Sorensen *et al.*, 1990). The difference in the translation rate of minor and major codons directly correlates with the content of the tRNA pool (Emilsson and Kurland, 1990). The more charged-tRNA available to translate minor codons, the faster these codons are translated. Generally the rate of translation of a particular codon doesn't have an overall effect on the translation rate of the mRNA or on the overall amount of protein produced. However, minor codons located near the start of a gene can reduce the amount of protein produced by preventing additional ribosomes from initiating translation (Sorensen *et al.*, 1990). This effect is most pronounced within the first 25 codons (the length covered by two ribosomes) (Chen and Inouye, 1990). Experiments have demonstrated that as the minor codons are moved further away from the start of the gene the effect is reduced (Chen and Inouye, 1994). Thus, according to this "minor codon modulator hypothesis" it is expected that the UUA codon, located at codon 32 of the *ccaR* transcript should have a moderate effect on the amount of protein translated in the *bldA* mutant. The results of this thesis work do not support this since approximately equivalent amounts of protein are observed in the wild-type and *bldA* mutant strains even though the amount of *ccaR* transcript may be reduced in the *bldA* mutant. There are several additional minor codons (codons used at a frequency of less than 5/1000 codons) located within the first 32 codons of the *ccaR* gene. These include ACA10, AAA13, GTA16, ATA18, CGA24, CGA27, and TCA31. The presence of these minor codons may reduce the initial translation rate of the *ccaR* transcript, which could have implications regarding the ability to mistranslate the TTA32 codon. However, this cannot be the only factor regulating mistranslation as the *bldA* dependent *strR* (streptomycin pathway-specific regulator) gene of *S. griseus* contains several minor codons within the first 30 codons (including TTA30). The location of the TTA codons in the *strR* and *ccaR* genes are similar, as well as the distribution of other minor codons, so the difference in *bldA* dependency cannot be accounted for solely by modulation of the translation rate.

In summary, the ability of *S. clavuligerus bldA* mutants to produce cephamycin C and clavulanic acid in the absence of tRNA $^{\text{Leu}}_{\text{UAA}}$ is due to the efficient mistranslation of the single UUA codon in the *ccaR* transcript, as the CcaR protein is required for antibiotic biosynthesis. The most likely candidate for mistranslation of UUA codons in *bldA* mutants is the $\text{tRNA}^{\text{Leu}}_{\text{CAA}}$, which differs from the $\text{tRNA}^{\text{Leu}}_{\text{UAA}}$ anticodon at position 34, the wobble position. Modifications to the $\text{tRNA}^{\text{Leu}}_{\text{CAA}}$, particularly at position 34 may enhance the wobble base pairing between C34 and the third codon position. Mistranslation by the $\text{tRNA}^{\text{Leu}}_{\text{CAA}}$ would allow the correct incorporation of the amino acid leucine at UUA codons. Although $\text{tRNA}^{\text{Leu}}_{\text{CAA}}$ seems the most likely candidate for mistranslation of UUA codons, other possible candidates such as $\text{tRNA}^{\text{Leu}}_{\text{UAG}}$, $\text{tRNA}^{\text{Phe}}_{\text{AAA}}$, or a suppressor tRNA such as the *supH* suppressor tRNA of *E. coli* ($\text{tRNA}^{\text{Ser}}_{\text{UAA}}$) (Leclerc *et al.*, 1989), should not be discounted without experimental evidence. However, in order to simplify further discussion, it will be assumed that the mistranslation of UUA codons in *bldA* mutants is by the $\text{tRNA}^{\text{Leu}}_{\text{CAA}}$.

Since it is apparent that the UUA codon in the *ccaR* transcript in the *S. clavuligerus bldA* mutant can be mistranslated and the *ccaR* gene is effectively *bldA* independent, it raises the question why aren't all UUA codons mistranslated in a *bldA* mutant and why aren't more genes *bldA* independent? There is evidence that suggests that the TTA codons in some of these *bldA* dependent genes are mistranslated. The regulatory gene for actinorhodin production, *act*II-ORF4, can restore actinorhodin production to a *bldA* strain if cloned in high copy number (Passantino *et al.*, 1991), although not to the levels observed in the wild-type strain (Fernández-Moreno *et al.*, 1991). This suggests that if the level of *act*II-orf4 transcript is increased there is a subsequent increase in the overall amount of functional protein produced by mistranslation of the single UUA codon, which is sufficient for activation of the actinorhodin biosynthetic pathway. The transcriptional activator of undecylprodigiosin, *redZ*, was first identified by virtue of its ability to produce red pigment in a *bldA* strain. One of the pigmented while bald (Pwb) mutants was found to carry a *redZ* up-promoter mutation, suggesting that an increase in the level of *redZ* transcription allowed sufficient mistranslation for the restoration of undecylprodigiosin production in a *bldA* mutant (Guthrie *et al.*, 1998). Multiple copies of *redZ* in a *bldA* mutant have a similar effect (White and Bibb, 1997). The reporter genes *hyg* and *aad* are only partially dependent on *bldA* for activity, and it appears that under some conditions there is a low level of mistranslation of the UUA codons, which yields enough product to confer antibiotic resistance (Leskiw *et al.*, 1991b). Taken together, the evidence indicates that most if not all UUA codons are mistranslated to some extent, however, except for the UUA codon in *ccaR* which is mistranslated efficiently, most UUA codons are mistranslated at low levels. So the question becomes, why is the UUA codon in *ccaR* the only one that is mistranslated efficiently?

Some of the factors that could affect the ability of *bldA* mutants to mistranslate UUA codons are environmental. Intracellular conditions are dictated by the metabolism of the organism and vary in response to a variety of signals such as temperature and nutritional status. All of these factors can affect the content of the tRNA pool, translation rates, and base modifications of tRNA molecules. Certainly, these factors must affect mistranslation rates, and the expression of RedZ in a *bldA* mutant under low phosphate conditions (White and Bibb, 1997), as well as the sporulation of *bldA* mutants on some media but not others, support this hypothesis. However, environmental conditions cannot be the only factors that dictate when a particular UUA codon is mistranslated in a *bldA* mutant, as *S. clavuligerus bldA* mutants efficiently mistranslate *ccaR* transcripts but not *gfp-mut3* transcripts under the same conditions. While it can be argued that depletion of $tRNA^{Leu}_{CAA}$ prevents the *gfp-mut3* transcript (which contains three UUA codons) from being translated, this possibility seems unlikely as the minor codon UUG is not nearly as rare as the UUA codon and it is expected that $tRNA^{Leu}_{CAA}$ should not become so easily limited. If the pool of $tRNA^{Leu}_{CAA}$ did become severely limited, a noticeable phenotype would be expected, as TTG codons are distributed randomly throughout the genome and thus occur in essential genes. Another argument is that ribosome stalling at UUA codons affects mRNA stability, so that while most targets are mistranslated, many aren't detected because the amount of protein produced is too low to detect. While most studies don't address the stability of a given mRNA molecule, they do look at the amount of a transcript present at a given time. These studies all demonstrate that the relative amounts of transcript do not differ significantly between wild-type and *bldA* mutant strains. If the stability of an mRNA was affected, then the amount of transcript detected by northern analysis would be expected to decrease. Therefore, taken together, all of the available information suggests that there is a context effect, something unusual about the context of the UUA codon in the *ccaR* transcript that allows it to be mistranslated efficiently.

To address the possibility that there is an unusual context effect in the *ccaR* transcript that favors its mistranslation, and makes it essentially *bldA* independent, the *ccaR* coding sequence was compared to all of the *bldA* dependent targets (including reporter genes from other species) to determine if there is something unique about the context of the TTA codon in the *ccaR* gene. The targets that have been shown to be *bldA* dependent are: *carB (S. thermotolerans), hyg (S. hygroscopicus), act*II-ORF2 *(S. coelicolor), act*II-ORF4 *(S. coelicolor), strR (S. griseus), redZ (S. coelicolor), celA2 (S. halstedii), lipR (S. exfoliatus* M11) (see Table I.1 for references); the non-streptomycete reporter genes *lacZ, aad, ampC* (Leskiw *et al.*, 1991b), and *gfp-mut3* (this work). The relative position of the TTA codons in these genes was analysed and the possible effect according to the "minor codon modulator hypothesis" was predicted (minor codons within the first 25 codons are predicted to have a severe effect on gene expression, moderate effect from 25-50 codons, and no effect past 50 codons). The presence of other minor codons immediately surrounding the TTA codon as well as the number of minor codons within 12 codons upstream (one ribosomal unit) was compared. The nucleotides surrounding the TTA codon were also compared. The results of the comparison are shown in Table IV.2.

The relative position of the TTA codon within the gene appears to have no effect on *bldA* dependency, suggesting that the modulation of expression by minor TTA codons has little effect, at least under conditions where the *bldA* tRNA is absent. The TTA codon in the *ccaR* gene is preceded by a minor codon (TCA) which could indicate that mistranslation occurs when the ribosome slows to translate the minor serine codon. However, the two TTA codons in the *carB* gene are also preceded by minor codons (CGA), as is the single TTA codon of the *strR* gene (AAA). So while the presence of a minor codon prior to the TTA codon may increase the rate of mistranslation of the TTA codon, it does not guarantee that it will be mistranslated. The presence or absence of a

Table IV.2 Comparison of *ccaR* (*bldA* independent target) to the *bldA* dependent targets

Independent target	Position of TTA (codon)	Modulator hypothesis prediction	Minor codon immediately upstream	Minor codon immediately downstream	# minor codons within 12 codons upstream	Nucleotide 5' of TTA	Nucleotide 3' of TTA
ccaR	32	moderate	Yes	No	3	A	G
Dependent targets*							
carB	95	none	Yes	No	2	A	T
	111	none	Yes	No	2	A	C
hyg	7	severe	No	Yes	2	G	C
act II-ORF2	19	severe	No	No	0	G	C
act II-ORF4	5	severe	No	Yes	1	C	T
strR (S. griseus)	30	moderate	Yes	No	2	A	T
redZ	156	none	No	No	1	C	C
celA2	4	severe	No	No	1	G	C
lipR	831	none	No	No	0	G	C
lacZ	12	severe	Yes	Yes	4	T	C
	178	none	Yes	No	4	T	C
	261	none	Yes	No	6	T	T
	563	none	No	No	3	T	C
	632	none	No	No	2	T	T
	850	none	No	Yes	3	C	T
	901	none	No	No	2	A	G
ampC	9	severe	No	Yes	1	C	T
	10	severe	Yes	Yes	2	A	A
	78	none	No	No	4	G	G
	101	none	No	No	4	G	A
	123	none	Yes	Yes	6	A	C
	254	none	Yes	Yes	4	T	A
aad	41	moderate	No	No	4	G	T
	58	none	No	No	5	G	A
	230	none	No	No	5	C	A
gfp-mut3	18	severe	No	Yes	7	A	G
	125	none	No	Yes	6	G	A
	199	none	Yes	Yes	8	T	C

* Refer to Table I.1 for references

minor codon immediately following the TTA codon has no apparent effect as the independent target *ccaR,* as well as the dependent targets *carB, act*II-ORF2, *strR, redZ, celA2, lipR,* and *aad* all lack a minor codon in this position while the dependent targets *hyg, act*II-ORF4, and *gfp-mut3* contain a minor codon in this position. The number of minor codons within one ribosomal unit of the TTA codon varies from zero (*act*II-ORF2 and *lipR)* to eight *(gfp-mut3)* among the dependent targets, which suggests that the translation rate in the vicinity of the TTA codon is not a factor in the mistranslation of TTA codons.

The nucleotide immediately 5' of the TTA codon is an adenine for the independent target *ccaR,* and for the dependent targets *carB* and *strR.* Of the ten TTA codons in the *Streptomyces* genes four of them are preceded by an adenine. This is somewhat unusual as this corresponds with the third codon position which typically shows a 90% bias of G+C in *Streptomyces*, however, a sample size of ten is not statistically significant, and as the adenine 5' of the TTA codon occurs in both dependent and independent targets it doesn't appear to be relevant.

The TTA codon in the *bldA* independent target *ccaR* is followed by a guanine, and while some of the dependent targets have this same TTAG pattern (*lacZ, ampC,* and *gfp-mut3)*, this pattern only occurs in genes that contain more than one TTA codon. Furthermore, the TTAG pattern in the dependent targets is only observed in the non-streptomycete reporter genes. It is possible that a TTAG sequence is translated more efficiently by $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{CAA}}$ as the guanine outside of the codon could stabilize the wobble base pairing between C34 and the third codon position. However, as discussed previously, the nucleoside 3' of a "hungry" codon can be an important determinant for frameshifting, particularly to the +1 frameshift event. For example, the codon AAG (lysine) sequence is "shifty" in *E. coli* when the amino acid lysine is depleted, but only when followed by a cytidine (and to a lesser extent uridine) residue. Although the situation with *bldA* differs in that there is a complete absence of $tRNA^{Leu}_{UAA}$ rather than depletion of the amino acid leucine resulting in an increase of uncharged $tRNA^{Leu}_{UAA}$, this could potentially have a similar effect as it also results in a shortage of charged, cognate-tRNA for a particular codon. If the effect is similar and the UUA codon is "shifty" in the absence of $tRNA^{Leu}_{UAA}$, then the nucleotide 3' of the UUA codon can specify whether a shift is permitted or whether the codon is mistranslated in-frame, and could be the difference between *bldA* dependent or independent targets.

It is becoming increasingly evident that there are two separate events that can occur at UUA codons, as TTA-containing genes can be either *bldA* dependent or independent. The first event occurs in *bldA* independent genes and is the efficient mistranslation of UUA codons in the absence of $tRNA^{Leu}_{UAA}$. At present, the only instance where this event appears to be prevalent is during translation of the single UUA codon in *ccaR*, and the G nucleotide 3' of the UUA codon appears to play a role in this efficient mistranslation. What occurs during the second event is not clear, however, it can be assumed that the UUA codons are not efficiently mistranslated, as functional product is not produced and the respective genes are *bldA* dependent. It has generally been assumed that in a *bldA* mutant the translation machinery stalls at UUA codons, and in most cases, the ribosomes eventually dissociate from the transcripts. While premature translation termination could be the second event, as it is consistent with the lack of functional target product, it seems unlikely that translational termination and dissociation of the ribosomal complex would be favored over mistranslation. A more likely possibility for this second event is that frameshifts are introduced at some UUA codons. It is not difficult to perceive how out-of-frame mistranslation may compete with in-frame mistranslation, as "hungry" codons can induce a variety of mistranslational events, and nucleotides 3' of the "hungry" codon can dictate whether a frameshift is allowed. A frameshift product would be neither functional nor detectable by western analysis, in most cases, which is consistent with the results to date. Thus, a new theory is presented to explain and predict what happens at UUA codons in *bldA* mutants.

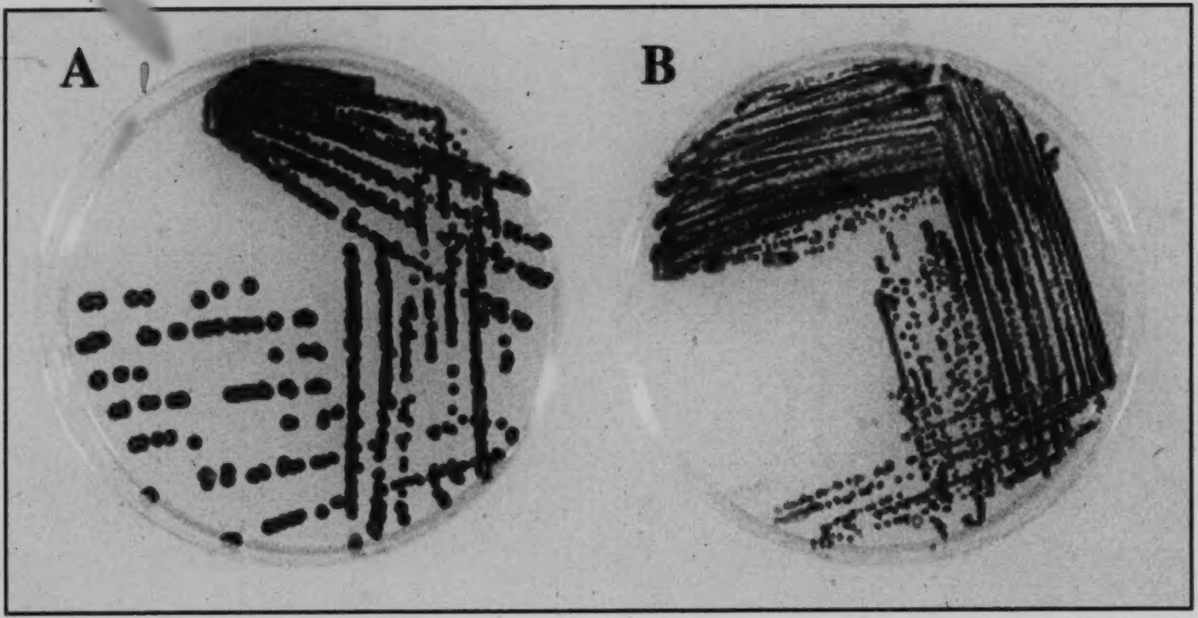
The premise of this new "mistranslation theory" is that all UUA codons are mistranslated in a *bldA* mutant, although there are two different mistranslation events that can occur. The two events are in-frame mistranslation and frameshifting. The frameshift event is favored over in-frame mistranslation unless the UUA codon is flanked on the 3' side by certain nucleotides which prevent the shift. The importance of the nucleotide 3' of the "hungry" UUA codon suggests that the frameshift event involves pairing of a "shifty" tRNA with the +1 frame codon, and thus introduces a +1 frameshift. Since the *bldA* target, *ccaR*, containing the sequence UUA G is *bldA* independent, this sequence must be mistranslated in-frame, and a G 3' of the UUA codon must prevent or suppress the +1 shift. The +1 reading frame in this sequence is the UAG amber stop codon, and since a +1 shift of this type requires a tRNA pairing with the +1 frame codon, a termination codon in the +1 frame should not allow a frameshift. Likewise, the sequence UUA A, which encodes an ochre stop codon in the +1 frame should not allow a +1 shift. So according to this new theory, the sequence UUA R (G or A) will be mistranslated exclusively in-frame in a *bldA* mutant. The sequence UUA Y (UUA C and UUA U), on the other hand, encodes tyrosine in the +1 frame, and will allow a +1 frameshift to occur. The frameshift event must occur preferentially over in-frame mistranslation as all *bldA* targets containing the sequences UUA Y are *bldA* dependent. Why the frameshift event is favored over in-frame mistranslation is not apparent, perhaps the tRNA$^{Tyr}_{GUA}$ that translates UAY codons is "shifty". It is also possible that tyrosinyl-tRNA is much more abundant than the leucyl-tRNA required for in-frame mistranslation and out-competes tRNA$^{Leu}_{CAA}$ for the ribosomal A site, or the codon-anticodon interaction is better with the tyrosinyl-tRNA in the +1 frame than it is with tRNA$^{Leu}_{CAA}$ in the zero frame. At any rate, the frameshift event is favored over in-frame translation, and translation at most transcripts containing the UUA Y sequence will shift to the +1 frame at the UUA Y sequence. Most of these frameshift products will be non-functional so their respective genes will be *bldA* dependent. However, it is possible for a gene containing a "shifty" TTA Y sequence to be *bldA* independent if the TTAY sequence is located 5' of an alternate zero frame initiation codon, or if the frameshift is introduced near the C-terminus, and there are no important structural or functional domains disrupted by the frameshift. However, the majority of the *bldA* independent genes will consist only of "non-shifty" TTA R sequences.

While UUA R codons are mistranslated exclusively in-frame as frameshifts can not be introduced at these sequences, it seems likely that UUA Y codons can be mistranslated either in-frame or out-of-frame, and the two events are likely competitive. While it is seems probable that UUA Y sequences are predominately translated out-of-frame, there is substantial evidence to support the premise that these UUA Y sequences are also mistranslated in-frame at some low level. In most instances, the level of in-frame mistranslation is too low to detect, as there is insufficient product to provide function or to detect by western analysis. However, increased levels of transcription of *act*II-orf4 or *redZ* can restore actinorhodin or undecylprodigiosin production to a *bldA* mutant, which indicates that there is some in-frame mistranslation of these two transcripts, and if the level of transcription is high enough these in-frame mistranslational products can be detected.

If these two events are competitive then it is likely that the ratio of frameshifting over in-frame mistranslation at TTA Y codons will vary depending on the intracellular conditions. For instance, if the ratio of $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{CAA}}$ to $\mathrm{tRNA}^{\mathrm{Tyr}}_{\mathrm{GUA}}$ increases significantly, the frequency of in-frame mistranslation should increase. Alternatively, if the ratio of $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{CAA}}$ to $\mathrm{tRNA}^{\mathrm{Tyr}}_{\mathrm{GUA}}$ decreases then the frequency of in-frame mistranslation should also decrease. Other factors that affect the translation rate, tRNA pool, or modification of tRNA bases can also either suppress or enhance one mistranslation event over the other. If sufficient product is produced, by in-frame mistranslation, under certain conditions, then the gene should appear to be *bldA* independent under those conditions. Partial *bldA* dependence is observed with both the *aad* and *hyg* reporter genes as they confer antibiotic resistance to *bldA* mutants when grown on the rich medium, R2YE, but not when grown on minimal medium (Leskiw *et al.*, 1991b). *S. coelicolor* and *S. lividans bldA* mutants are able to sporulate on minimal media containing mannitol but are unable to sporulate on glucose containing media (Champness, 1988; Merrick, 1976), and *S. griseus bldA* mutants sporulate on minimal media but not on nutritionally rich media (McCue *et al.*, 1992). Red pigment (undecylprodigiosin) production can be restored to a *bldA* mutant when grown on media containing low concentrations of phosphate (Figure IV.3), and

Figure IV.3

Production of undecylprodigiosin (red pigment) by a *bldA* mutant. Panel A shows red pigment production by *S. coelicolor* J681(*ΔbldA)* on minimal medium + mannitol. Panel B shows red pigment production by the closely related *S. lividans* J1725 *(bldA-)* on the same medium.


A
B

redZ and *act*II-orf4, in high copy number only restore actinorhodin and undecylprodigiosin production to *bldA* mutants grown on rich medium but not the relatively simple and defined SMMS (White and Bibb, 1997). Taken together, all of these results suggest that the level of in-frame mistranslation of TTA Y codons varies depending on intracellular conditions, however, the evidence is so far only circumstantial and additional definitive tests that specifically address this issue are required.

Although the main consequence of frameshifting will be the production of non-functional products, the pool of charged tRNA may also be affected. Since *Streptomyces* genes are highly biased for the use of G+C in the third codon position (90%) and first codon position (70%), changing the reading frame to the +1 frame changes the bias so the third codon position is only 70% G+C, and the 20% increase in A+T bases at the third codon position will presumably generate a corresponding increase in the number of minor codons specified. If there is a UUA Y sequence in the +1 frame, a second +1 shift can occur which will increase the number of minor codons further as the third codon position will fall to 50% G+C. Although it is also expected that there will also be a corresponding increase in the number of stop codons encoded in the +1 and +2 frames, and most polypeptides will be terminated shortly after shifting to the +1 frame, there may some instances where a substantial frameshift extension product is generated. If there is sufficient out-of-frame translation, depletion of tRNAs that decode minor codons could potentially occur. The content of amino acids incorporated into polypeptide will also be affected by out-of-frame translation. A +1 frameshift changes the first codon position to 50% G+C. The 20% decrease in G+C content decreases the number of codons specifying amino acids such as leucine, proline, alanine, arginine, and valine and increases the number of codons specifying amino acids such as serine, tyrosine, cysteine, tryptophan, and threonine. This could potentially lead to starvation conditions for infrequently used amino acids such as, cysteine and tryptophan, as well as depletion of the tRNAs that decode the minor serine and threonine codons. Although it is not known what effect depletion of minor tRNAs or starvation for minor amino acids will have on mistranslation events, it is possible that under these conditions the stringent response will be induced which could globally alter gene expression at the transcriptional level.

One of the features of this "mistranslation theory" is that *bldA* dependence or independence of target genes can be predicted by analyzing the location and context of the TTA codons. All of the *Streptomyces bldA* targets have been compiled and separated into two main classes (Table IV.3). The various reporter genes that have been used to assess translation of TTA codons in *bldA* mutants are shown in Table IV.4. Class I genes, containing TTA R sequences, are expected to be translationally independent of *bldA*. Class II genes are predicted to be translationally dependent on *bldA* for expression. As outlined above, TTA R sequences are predicted to be exclusively mistranslated in-frame, and should be *bldA* independent. If this is true, there are twelve targets, in addition to *ccaR*, that should be translationally independent of *bldA*. Since the *bldA* gene has not been identified in most of the species that contain the predicted *bldA* independent targets it is not surprising that they have yet to be tested for *bldA* dependence. The *bldA* gene, however, has been identified in *S. griseus*, so tests to confirm that the *bldA* target *attL* is *bldA* independent are warranted. One of the Class I targets, *strN* from *S. glaucescens*, is expected to be transcriptionally dependent on *bldA*, as it is regulated by *strR*, a target predicted to be *bldA* dependent, although translation of *strN* is expected to

Table IV.3 Classification of *Streptomyces* genes containing TTA codons

Gene	Organism	# TTA codons	Position° (amino acid)	bldA dependence (experimental)	Target sequence	bldA dependence (predicted)	Comments
Class I: *bldA* independent							
ccaR	*S. clavuligerus*	1	38 /262	not dependent	TTAG	not dependent	Western analysis and antibiotic production confirms CcaR production in *bldA* mutant
pur6	*S. anulatus*	1	29/ 338	N/D*	TTAA	not dependent	
ORF7	*S. bambergiensis*	1	3 /109	N/D	TTAG	not dependent	
tlrD	*S. fradiae*	1	76 /327	N/D	TTAG	not dependent	
tylG- ORF3	*S. fradiae*	1	1271 /3729	N/D	TTAG	not dependent	
strN	*S. glaucescens*	1	2 /315	N/D	TTAG	not dependent	Translationally *bldA* independent, transcriptionally *bldA* dependent (StrR regulated)
attL	*S. griseus*	1	635/ 666	N/D	TTAA	not dependent	
kmr	*S. kanamyceticus*	1	12 /276	N/D	TTAA	not dependent	
mdmB	*S. mycarofaciens*	1	20 /387	N/D	TTAG	not dependent	
SMPI	*S. nigrescens*	1	7 /102	N/D	TTAG	not dependent	TTA codon is in the signal peptide
fkbA	*Streptomyces sp.*	1	4185 /6420	N/D	TTAG	not dependent	
acyB2	*S. thermotolerans*	1	228 /387	N/D	TTAA	not dependent	
secE	*S. virginiae*	1	7 /121	N/D	TTAG	not dependent	
Class II: *bldA* dependent							
cmcT	*S. clavuligerus*	1	470/523	N/D	TTAC	dependent	Dependent if C-terminal region essential
act II-ORF2	*S. coelicolor*	1	19 /578	dependent	TTAC	dependent	Not known if regulated by *act* IIORF4 (*bldA* dependent target)
act II-ORF4	*S. coelicolor*	1	5 /255	dependent	TTAT	dependent	*actII* ORF4 in high copy number restores actinorhodin production to a *bldA* mutant on rich medium[2]
redZ	*S. coelicolor*	1	156 /217	dependent	TTAC	dependent	*redZ* in high copy number restores red pigment production to a *bldA* mutant on rich medium, red productin also restored under low [phosphate], *redZ* autoregulated
bldX (*nrsA*)	*S. coelicolor*	1	207 /473	N/D	TTAC	dependent	*bldX* encodes two proteins from a single ORF, in high copy number can restore sporulation to *S. griseus bldA* mutants
nrsA (ORF1590)	*S. griseus*	1	263 /529	N/D	TTAC	dependent	*nrsA* encodes two proteins from a single ORF, in high copy number can restore sporulation to *bldA* and A-factor mutants
strR	*S. griseus*	1	30 /350	dependent	TTAT	dependent	Regulator of *str* cluster (bldA mutants str-)
strN	*S. griseus*	2	2 /339 36/339	N/D	TTAG TTAC	independent target dependent	First TTA codon predicted to be *bldA* independent, *strN* regulated by StrR (*bldA* regulated)

Gene	Organism	# TTA codons	Position (amino acid)	bldA dependence (experimental)	Target sequence	bldA dependence (predicted)	Comments
Class II: *bldA* dependent continued							
orf	*S. griseus*	1	47 /297	N/D	TTAC	dependent	Not known if regulated by StrR
amfR	*S. griseus*	1	141 /201	N/D	TTAT	dependent	*amfR* in high copy number can restore sporulation to A-factor mutants
strR	*S. glaucescens*	1	30 /424	N/D	TTAT	dependent	Regulator of *str* cluster, *bldA* gene not yet identified in *S. glaucescens*
strA (sph)	*S. glaucescens*	1	39 /307	N/D	TTAC	dependent	Transcription also dependent on *bldA* (StrR regulated)
strD	*S. glaucescens*	1	14 /356	N/D	TTAC	dependent	Transcription also dependent on *bldA* (StrR regulated)
strS	*S. glaucescens*	1	7 /378	N/D	TTAC	dependent	Transcription also dependent on *bldA* (StrR regulated)
strV	*S. glaucescens*	1	4 /584	N/D	TTAC	dependent	Transcription also dependent on *bldA* (StrR regulated)
tcmP	*S. glaucescens*	1	226 /271	partially	TTAC	dependent	Dependent if C-terminal region essential
lipR	*S. exfoliatus* M11	1	831 /934	dependent	TTAC	dependent	Tested in *S. coelicolor bldA* mutant, TTA→CTC correction alleviated *bldA* dependence
celA2	*S. halstedii*	1	4 /377	dependent	TTAC	dependent	Tested in *S. coelicolor bldA* mutant, translation of downstream ORF also affected
hyg	*S. hygroscopicus*	1	7 /332	dependent	TTAC	dependent	Tested in *S. lividans* and *S. coelicolor bldA* mutants, (partial dependence), low levels of *hyg* gene product confer resistance[1]
celB	*S. lividans*	2	4 /381	N/D	TTAC	dependent	Correction of TTA codons decreases expression due to formation of secondary structures in the signal peptide region[4]
			22/381		TTAG	independent target	
carB	*S. thermotolerans*	2	95 /283	dependent	TTAT	dependent	Tested in *S. lividans* and *S. coelicolor bldA* mutants, both targets predicted to be *bldA* dependent[1]
			111/283		TTAC		
srmR	*S. ambofaciens*	1	237 /604	N/D	TTAC	dependent	Mutations downstream of the TTA codon within the ORF disrupt function
traSA	*S. ambofaciens*	1	46 /415	N/D	TTAC	dependent	
orf183	*S. ambofaciens*	1	12 /183	N/D	TTAC	dependent	
OLE-ORF3	*S. antibioticus*	2	242 /3519	N/D	TTAC	dependent	
			770 /3519		TTAC		
abaB	*S. antibioticus*	1	122 /301	N/D	TTAC	dependent	
pur10	*S. anulatus*	1	29 /338	N/D	TTAC	dependent	
mtmR	*S. argillaceus*	2	7 /276	N/D	TTAC	dependent	Although second TTA codon should not shift +1, other potential shifts are possible
			227 /276		AATTTAA	independent target	

Gene	Organism	# TTA codons	Position (amino acid)	bldA dependence (experimental)	Target sequence	bldA dependence (predicted)	Comments
Class II: *bldA* dependent continued							
rnaSa3	*S. aureofaciens*	1	52 /181	N/D	TTAC	dependent	
ORF4	*S. bambergiensis*	2	80 /98	N/D	TTAC	dependent	
			83 /98		TTAC		
nidA1	*S. caelestis*	1	4004 /4340	N/D	TTAC	dependent	
nidA4	*S. caelestis*	1	155 /1569	N/D	TTAC	dependent	
ORF	*S. carbophilus*	1	67 /410	N/D	TTAC	dependent	
SC10A5.18	*S. coelicolor*	1	4 /411	N/D	TTAC	dependent	
SCD78.22	*S. coelicolor*	1	33/339	N/D	TTAC	dependent	
SCI35.14	*S. coelicolor*	1	167 /180	N/D	TTAC	dependent	MutT signature is upstream from TTA codon, *bldA* dependent if C-term region essential
SCI35.26	*S. coelicolor*	1	44 /175	N/D	TTAC	dependent	
SC1C3.22	*S. coelicolor*	1	56 /442	N/D	TTAC	dependent	
SC1F2.20	*S. coelicolor*	2	17 /731	N/D	TTAG	independent target	Putative ATP/GTP binding domain downstream from second TTA codon
			184/731		TTAC	dependent	
SC3C8.03	*S. coelicolor*	1	249 /589	N/D	TTTAC	dependent	-1 or +1 shift possible
SC3C8.05	*S. coelicolor*	1	7 /127	N/D	TTAC	dependent	
SC3C8.06	*S. coelicolor*	1	45 /188	N/D	TTAC	dependent	
SC3C8.21	*S. coelicolor*	1	44 /175	N/D	TTAC	dependent	
SC4G2.12	*S. coelicolor*	1	21 /605	N/D	TTAC	dependent	
SC4H2.07	*S. coelicolor*	1	8 /400	N/D	TTAC	dependent	
SCRH2.20	*S. coelicolor*	1	525 /532	N/D	TTAC	dependent	Dependent if C-term region essential, frameshift could add C-terminal extension
SC6A9.34	*S. coelicolor*	1	82 /76[illegible]	N/D	TTAC	dependent	
SC8A6.16	*S. coelicolor*	1	17 /380	N/D	TTAC	dependent	
SC9C7.12	*S. coelicolor*	1	12 /197	N/D	TTAC	dependent	
spcR	*S. flavopersicus*	1	16 /330	N/D	TTAC	dependent	
tyl- ORF2	*S. fradiae*	1	264 /452	N/D	TTAC	dependent	
tylG- ORF1	*S. fradiae*	1	3521 /4472	N/D	TTAC	dependent	
tylG- ORF2	*S. fradiae*	1	1263 /1864	N/D	TTAC	dependent	
axnA	*S. globisporus*	1	3 /143	N/D	TTAC	dependent	TTA codon is in the leader protein sequence
pkg4	*S. granaticolor*	1	114 /761	N/D	TTAC	dependent	
suaC	*S. griseolus*	1	26 /406	N/D	TTAC	dependent	
secY	*S. griseus*	1	248 /436	N/D	TTAC	dependent	
brpA	*S. hygroscopicus*	1	250 /256	N/D	TTAC	dependent	Dependent if C-term region essential
rapH	*S. hygroscopicus*	1	38 /872	N/D	TTAC	dependent	TTA codon is upstream from the putatitive helix-turn-helix motif
rapG	*S. hygroscopicus*	1	252 /330	N/D	TTAC	dependent	TTA codon is within the putatitive helix-turn-helix motif
rapB	*S. hygroscopicus*	1	4578 /10223	N/D	TTAC	dependent	

Gene	Organism	# TTA codons	Position (amino acid)	bldA dependence (experimental)	Target sequence	bldA dependence (predicted)	Comments
Class II: ***bldA*** **dependent continued**							
rapP	*S. hygroscopicus*	1	768 /1541	N/D	TTAC	dependent	
orf4	*S. hygroscopicus*	1	655 /1937	N/D	TTAC	dependent	
orf6	*S. hygroscopicus*	1	369 /948	N/D	TTAC	dependent	
pcbAB	*S. lactamdurans*	2	775 /3649	N/D	TTAC	dependent	
			3456 /3649		TTAC		
TYR1-ORF1	*S. lavendulae*	1	4 /156	N/D	TTAT	dependent	
mcr- ORF3	*S. lavendulae*	1	83 /281	N/D	TTAC	dependent	
lmbB2	*S. lincolnesis*	1	104 /317	N/D	TTAC	dependent	
lmbY	*S. lincolnesis*	1	19 /295	N/D	AATTAT	dependent	Other shifts possible
lmbU	*S. lincolnesis*	1	6 /223	N/D	AATTTAT	dependent	Other shifts possible
aac(3) -Ia	*S. lividans*	2	2 /177	N/D	TTAC	dependent	The first TTA codon will induce a +1 frameshift which will terminate at the TTAG sequence
			39 /177		TTAG	independent target	
amy	*S. lividans*	1	731 /919	N/D	TTAC	dependent	
mdmA	*S. mycarofaciens*	3	59 /310+	N/D	TTAG	independent target	TTTTAA could potentially cause -1 or -2 shift, +1 shift not allowed, TTTAC could potentially cause -1 or +1 shift
			63/310+		TTTTAA	unclear	
			85/310+		TTTAC	dependent	
snoA	*S. nogalater*	1	76 /665	N/D	AATTAT	dependent	
snoG	*S. nogalater*	1	328 /328	N/D	TTAT	dependent	+1 frameshift will generate a significantly longer peptide sequence
dnrO	*S. peucetius*	1	86 /340	N/D	TTAC	dependent	
chtA	*S. plicatus*	1	276 /610	N/D	TTAC	dependent	
eglS	*S. rochei*	1	4 /382	N/D	TTAC	dependent	
dauH	*Streptomyces sp* C5	1	269 /442	N/D	TTAC	dependent	
ORF	*Streptomyces sp* TH1	1	36 870	N/D	TTAC	dependent	
choP	*Streptomyces* SA-COO	1	56 /381	N/D	TTAT	dependent	
pabAB	*S. venezuelae*	1	1779 /1949	N/D	TTAC	dependent	
phsA	*S. viridochromogenes*	1	222 /622	N/D	TTAC	dependent	

*N/D - not determined

+Position of TTA codon/Total number of codons

References are the same as indicated in Table I.1 with the exception of the following:

1. Leskiw *et al* . (1991)
2. Passantino *et al.* (1991), Gramajo *et al.* (1993)
3. Ueda *et al.* (1993)
4. S. Wittman, personal communication

Table IV.4 Comparison of codon context in TTA-containing reporter genes

Gene	Organism	# TTA codons	Position (amino acid)	bldA dependence (experimental)	Target sequence	bldA dependence (predicted)	Comments
Reporter genes containing TTA codons							
carB [1]	*S. thermotolerans*	2	95 /283	dependent	TTAT	dependent	Tested in *S. lividans* and *S. coelicolor bldA* mutants, both targets predicted to be *bldA* dependent
			111/283		TTAC		
hyg	*S. hygroscopicus*	1	7 /332	dependent	TTAC	dependent	Tested in *S. lividans* and *S. coelicolor bldA* mutants, (partial
mut3gfp [2]	*Aequorea victoria*	3	18/238	dependent	TTAG	independent target	First two TTA codons predicted to be *bldA* independent, third TTA codon could have alternate shifts
			125/238		TTAA	independent target	
			199/238		TTTTAC	dependent	
aad [1]	*E. coli* plasmid R538-1	3	41/323	dependent	TTAT	dependent	Tested in *S. lividans* and *S. coelicolor bldA* mutants, (partial dependence), low levels of aad gene product confer resistance
			58/323		TTAA	independent target	
			230/323		TTAA	independent target	
ampC [1]	*E. coli*	6	9/377	dependent	TTAT	dependent	Severe *bldA* dependence, predicted that TTA at position 10, 78, 101, and possibly 254 are *bldA* independent, TTTAAAA target could potentially shift -1
			10/377		TTAA	independent target	
			78/377		TTAG	independent target	
			101/377		TTAA	independent target	
			123/377		TTAC	dependent	
			254/377		TTTAAAA	unclear	
lacZ [1]	*E. coli*	7	12/1024	dependent	TTAC	dependent	Severe *bldA* dependence, only the TTA at postion 901 is expected to be *bldA* independent
			178/1024		TTAC	dependent	
			261/1024		TTAT	dependent	
			563/1024		TTAC	dependent	
			632/1024		TTAT	dependent	
			850/1024		TTAT	dependent	
			901/1024		TTAG	independent target	
ssi [3]	*S. albogriseolus* S-3253	0→2	79/113	dependent	CTG→TTAT	dependent	CTG-CTC to TTA-TTA reduced the production of SSI markedly in *S. lividans*
			80/113		CTC→TTAA	independent target	

1. Leskiw et al (1991)
2. Cormack et al (1996), Markus and Leskiw (1997)
3. Ueda et al (1993)

be *bldA* independent. The remainder of *bldA* targets are predicted to be translationally *bldA* dependent (Class II), although several of them, with more than one TTA codon contain predicted *bldA* independent targets TTA R. There are a few cases where the TTA R codon is in a context that may allow alternate shifts. For example, codon 63 of *mdmA* of *S. mycarofaciens* contains the sequence TT TTA A, and while this sequence will not induce a shift to the +1 frame at the TAA sequence, it could potentially induce a –1, or -2 shift at the TTTT sequence. Likewise there are repetitive sequences around TTA Y "shifty" codons, such as the TT TTA C sequence around codon 199 of *gfp-mut3*. This sequence could potentially shift –2 or –1 at the TTTT sequence, or shift +1 at the TAC sequence. It would be interesting to determine if repetitive sequences around TTA codons do in fact induce alternate frameshifts in *bldA* mutants, as this may indicate whether the "hungry" TTA codon itself is inherently "shifty" or if only the $\text{tRNA}^{\text{Tyr}}_{\text{GUA}}$ that introduces +1 frameshifts is "shifty".

One target of note is *cmcT*, which is the second *bldA* target to be identified in *S. clavuligerus*. The *cmcT* gene is a Class II target, so it is expected to be *bldA* dependent as the TTA C sequence is expected to induce a +1 frameshift. The gene maps within the cephamycin C biosynthetic gene cluster (see Figure III.9.3) and is predicted to encode an export gene for cephamycin C and/or clavulanic acid (Martin, 1998). Since the bioassays for antibiotic production in the *bldA* mutant were performed with either supernatant from liquid cultures or the surrounding medium from surface-grown cultures, it is apparent that these antibiotics were exported, which seems to contradict the prediction that *cmcT* is *bldA* dependent. However, a *cmcT* disruption reduced, but did not eliminate cephamycin C export (S. Jensen, personal communication), indicating that there are additional elements involved in the export of cephamycin C. While cephamycin C production did not appear to be reduced in the *bldA* mutant, the appearance of this antibiotic did seem somewhat delayed. Perhaps the delayed appearance is due to less efficient export and not due to delayed transcription of the genes in the biosynthetic pathway or an overall slower growth rate. Another factor that may affect the *bldA* dependence of *cmcT* is the location of the TTA codon near the C-terminus of the gene (codon 470 of the 523 amino acid protein). The frameshift product could potentially be partially or fully functional, depending on the importance of the C-terminus. Also, there may be significant levels of in-frame mistranslation of the *cmcT* transcript. The *tcmP* gene of *S. glaucescens,* which contains a "shifty" TTA codon near the C-terminus (226/271) produces approximately 1/6 the wild-type level of functional protein in a *bldA* mutant. It seems likely that the TTA codon in this target is mistranslated in-frame at significant levels. While the reason for the apparently high level of in-frame mistranslation is not clear, one possibility is that in-frame mistranslation rates are higher when the TTA codon is far removed from the 5' end of the transcript. If this is the case then *cmcT* may also be mistranslated in-frame at significant levels. At any rate, further testing will have to be performed to determine what effect *bldA* has on *cmcT* expression.

There appears to be a preference for a cytidine following TTA codons as 71% (76/107) of the TTA codons are in the context TTA C. The nucleotide 3' of the TTA codons corresponds to the first codon position which is expected to show a G+C bias of 70%, however, the G+C content of nucleotides immediately 3' of TTA codons is 84% (76+14/107)with guanine nucleotides making up only 13 % (14/107) of this sum. Since protein sequences are not random, the apparent preference for cytidine following the

TTA codons could represent a preference for the amino acids leucine, proline, histidine, glutamine, and arginine following leucine residues and not a preference per se for the cytidine residue. If there is a preference for one or more amino acid(s) following leucine rather than a preference for a cytidine 3' of TTA codons then a similar distribution of nucleotides should be observed following all leucine codons. To address the possibility that there is a preference for one or more amino acids following leucine codons, the nucleotides 3' of leucine codons were assessed in 52 *bldA* targets (19 817 codons) (Table IV.5). In all, there were 2 202 leucine codons examined (58 TTA, 15 CTA, 76 TTG, 1186 CTG, 62 CTT, and 805 CTC). While there appears to be a preference for a cytidine nucleotide following the minor leucine codons TTA (70.7 %), CTA (46.7 %), TTG (40.8 %), and CTT (61.3 %), overall there is a slight preference for a guanine nucleotide following a leucine codon (40.3 % G versus 34.2% C). This indicates that the preference for cytidine nucleotides following TTA codons does not reflect a preference for a particular amino acid. Therefore, there is a preference in *Streptomyces* genes for the "shifty" TTA C sequence over the "non-shifty" TTA G sequence. In addition, there appears to be a preference for the "shifty" TTA T sequence over the "non-shifty" TTA A sequence as the distribution of thymidine nucleotides following TTA codons is 12.1 % whereas it is only 8.6% following all leucine codons.

The preference for "shifty" TTA Y sequences is similar to the preference in *E. coli* for the "shifty" AAG C sequence (Shpaer, 1986), which suggests that frameshifting may be a form of translational regulation. "Shifty" sequences may function as sensors of environmental conditions and ensure that the appropriate amount of functional product is produced under a variety of conditions. For example, in

Table IV.5 Distribution of nucleotides following leucine codons in *Streptomyces* genes containing TTA codons

Codon	Number	Percentage
TTA A	1	1.7
TTA T	7	12.1
TTA G	9	15.5
TTA C	41	70.7
Total	58	
% G+C		86.2

Codon	Number	Percentage
CTA A	3	20.0
CTA T	3	20.0
CTA G	2	13.3
CTA C	7	46.7
Total	15	
% G+C		60.0

Codon	Number	Percentage
TTG A	12	15.8
TTG T	11	14.5
TTG G	22	28.9
TTG C	31	40.8
Total	76	
% G+C		69.7

Codon	Number	Percentage
CTG A	171	14.4
CTG T	97	8.2
CTG G	436	36.8
CTG C	482	40.6
Total	1186	
% G+C		77.4

Codon	Number	Percentage
CTT A	3	4.8
CTT T	10	16.1
CTT G	11	17.7
CTT C	38	61.3
Total	62	
% G+C		79.0

Codon	Number	Percentage
CTC A	179	22.2
CTC T	62	7.7
CTC G	409	50.8
CTC C	155	19.3
Total	805	
% G+C		70.0

Codon	Number	Percentage
Leu A	369	16.8
Leu T	190	8.6
Leu G	889	40.3
Leu C	754	34.2
Total	2202	
% G+C		75.6

The results are compiled from the analysis of 52 TTA-containing genes representing 19,817 codons. The frequency of TTG (3.83/1000), CTG (59.85/1000), CTT (2.93/1000), and CTC (40.62/1000) codons is similar to the frequencies tabulated in Streptomyces codon-usage tables, indicating that the distribution of these codons is not biased in TTA-containing genes. The frequency of CTA codons is 0.76/1000 codons, which is significantly higher than the frequency of 0.4 tabulated for all genes, the bias likely reflects a preference for rare codons in minor genes (or preference against rare codons in major genes).

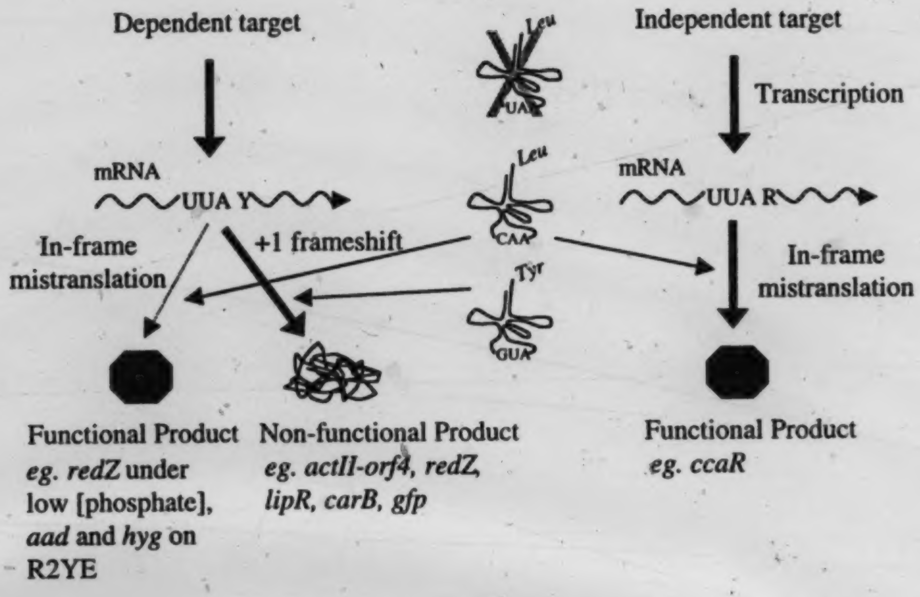
Streptomyces, if the growth conditions are sub-optimal for sporulation and antibiotic production, frameshifting at TTA Y sequences may reduce the amount of functional product produced to activate the various pathways involved in antibiotic production and sporulation. This would ensure that the amount of transcription of the pathway genes (and ultimately the amount of antibiotics produced or amount of spores) is appropriate for the intracellular conditions. The level of frameshifting could be regulated by the ratio of $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{UAA}}$ to $\mathrm{tRNA}^{\mathrm{Tyr}}_{\mathrm{GUA}}$, and either an increase in the pool of $\mathrm{tRNA}^{\mathrm{Tyr}}_{\mathrm{GUA}}$ or decrease in $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{UAA}}$ pool would increase the frequency of the frameshift event. Since accumulation of the *bldA* tRNA is temporally regulated, frameshifting could also be used as a means of preventing or reducing premature translation of *bldA* targets (Figure IV.4). At present, the role of the *bldA* tRNA and frameshifting in the modulation of gene expression is merely speculation, as it has yet to be proven that frameshifting occurs in *bldA* mutants, let alone in wild-type systems. However, it is well established that antibiotic production and sporulation are enhanced or reduced depending of the growth medium, and this is perhaps in part due to variations in the ratio of in-frame translation to frameshifting at *bldA* targets. Certainly, future experiments should investigate whether frameshifting at *bldA* targets occurs in wild-type strains, and what effect growth media composition has on the level of frameshifting.

While a role has been proposed for "shifty" TTA codons, the same cannot be said for "non-shifty" TTA codons, as these targets are believed to be translated in-frame under all conditions. Of the *bldA* targets, 13.5 % (13/96), including *ccaR*, contain only the "non-shifty" TTA R codons, and their expression is presumably *bldA* independent. These targets may have become independent of *bldA* for expression by selection or random

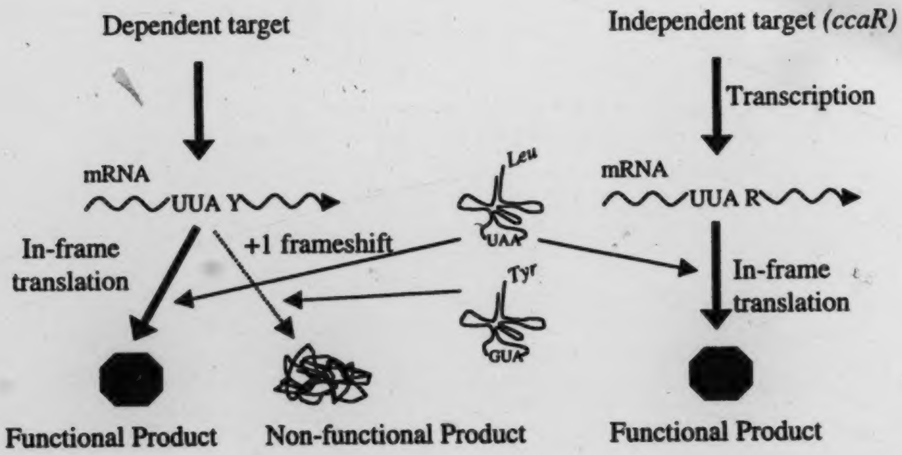
Figure IV.4

Translation of UUA codons in *Streptomyces* according to the "mistranslation theory". **(A)** *bldA* mutant. In the absence of $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{UAA}}$, the $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{CAA}}$ mistranslates UUA R codons exclusively in-frame to produce functional product, as is the case with *ccaR*. UUA Y codons are mistranslated in-frame by $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{CAA}}$ and out-of-frame by $\mathrm{tRNA}^{\mathrm{Tyr}}_{\mathrm{GUA}}$ to produce functional and non-functional products respectively. The +1 frameshift event is favored over in-frame mistranslation, so that most UUA Y containing genes appear *bldA* dependent. Environmental conditions, mediated by the ratio of $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{CAA}}$ to $\mathrm{tRNA}^{\mathrm{Tyr}}_{\mathrm{GUA}}$, can increase the frequency of in-frame mistranslation to produce sufficient functional product to make the target gene *bldA* independent, as is the case with *redZ* under low phosphate conditions, and *aad* and *hyg* on rich (R2YE) medium. **(B)** wild-type. Although most UUA codons are translated in-frame by $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{UAA}}$, there is a low frequency of +1 frameshifting by $\mathrm{tRNA}^{\mathrm{Tyr}}_{\mathrm{GUA}}$ at UUA Y codons. Frameshifting is used to modulate the expression of UUA Y containing genes so that the appropriate amount of functional product is produced under a given set of conditions.

A) ***Streptomyces bldA*** **mutant**


Dependent target
mRNA
UUA Y
In-frame mistranslation
+1 frameshift
Leu
UAA
Leu
CAA
Tyr
GUA
Independent target
Transcription
mRNA
UUA R
In-frame mistranslation
Functional Product *eg. redZ* under low [phosphate], *aad* and *hyg* on R2YE
Non-functional Product *eg. actII-orf4, redZ, lipR, carB, gfp*
Functional Product *eg. ccaR*

B) ***Steptomyces*** **wild-type strain**


Dependent target
mRNA
UUA Y
In-frame translation
+1 frameshift
Leu
UAA
Tyr
GUA
Independent target *(ccaR)*
Transcription
mRNA
UUA R
In-frame translation
Functional Product
Non-functional Product
Functional Product

mutation; alternately, they may still be evolving toward *bldA* dependence. Whatever the cause, the existence of *bldA* independent targets suggest that the modulation of expression by *bldA* is not essential for survival. This does not exclude the premise that in most cases TTA codons do function to optimize expression of a particular subset of genes for a given set of conditions; it merely indicates that most (if not all) *bldA* targets are highly regulated at the transcriptional level so that translational regulation is not essential.

While the "mistranslation translation" theory can explain why some *bldA* targets are independent while others are dependent, and can be expanded to provide a role for TTA codons in modulation of gene expression in wild-type systems, there is as yet no direct evidence to support or disprove this theory. Since the key element of this theory is that TTA Y codons induce frameshifts, and TTA R sequences do not, future experiments should address this issue. Frameshift products can be easily detected experimentally by placing reporter genes downstream of the expected shift sites in contexts where frameshifting is required for expression. For example the *act*II-ORF4::*ermE* fusion experiments (Gramajo *et al.*, 1993) should be attempted with the *ermE* gene in the +1 frame. The TTA T "shifty" codon could be changed to the TTA R "non-shifty" sequence with the *ermE* gene in the zero and +1 frame to demonstrate that TTA R sequences are mistranslated in-frame. The *act*II-ORF4 gene is a good candidate for translational frameshift fusion studies with other reporter genes because the TTA codon is near the start of the gene and therefore translational fusions would only have small N-terminal extensions on the reporter gene. This increases the likelihood that the translational fusion reporter gene would be functional, and could allow the use of a reporter gene such as *xylE* which can be assayed quantitatively. Other genes could certainly be tested with similar fusion studies, however, the fusion protein might only be detectable by western analysis, making quantitation more difficult. The frameshift event should also be tested on both the complex medium R2YE and on the relatively simple medium, SMMS, as the level of in-frame mistranslation is expected to be higher on R2YE medium, as both *act*II-orf4 and *redZ*, in high copy number, activate their respective pathways in a *bldA* mutant when grown on R2YE but not SMMS (White and Bibb, 1997). Equally important would be sensitive tests to determine the level of in-frame mistranslation under various conditions. Replacing the *act*II-ORF4 promoter with the constitutively expressed *ermE** promoter in in-frame *act*II::*ermE* fusion constructs should allow detection of the *ermE* product by western analysis. The levels of product could then be compared semi-quantitatively with the amount of product produced in the wild-type strain as well as under a variety of metabolic conditions.

While out-of-frame fusions are the best experiments to provide conclusive proof that frameshifts occur, other experiments could also provide support for this theory. According to predictions of the theory, the *gfp-mut3* gene has only one "shifty" TTA codon. Replacing the TTA199 with the CTA or TTG leucine codon should result in detectable GFP in a *bldA* mutant, whereas replacement of either of the other two TTA codons is expected to have no effect. Likewise, the *redZ* protein should be *bldA* independent under all conditions if the TTA CTG is changed to TTA GTG. This would change the codon 3' to the TTA codon to a valine instead of a leucine and would have to be tested in a wild-type strain to ensure that this amino acid substitution doesn't impair RedZ function. Additional experiments could test the translation of the *ccaR* transcript in a *S. coelicolor* or *S. lividans bldA* mutant, to demonstrate that the *bldA* independence of *ccaR* is due to context and not environment. This could also be demonstrated by changing the TTA GCG sequence to TTA CCG, a "shifty" sequence in the *S. clavuligerus bldA* strain. Again, this would change the amino acid following the leucine codon from an alanine to a proline and would require testing in the wild-type strain to determine if the substitution generates a functional protein. An alternate experiment to make *ccaR bldA* dependent would be to change a CTC, CTG, or TTG leucine codon to a TTA Y "shifty" leucine codon. The level of frameshifting in *bldA* strains should also be reduced by cloning additional copies of the $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{CAA}}$ gene. If the level of frameshifting does decrease, this would indicate that this is the tRNA responsible for mistranslation of UUA codons. If it doesn't decrease the level of frameshifting, then mistranslation by $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{UAG}}$ should be investigated as it is also a potential candidate for mistranslation of UUA codons.

If the experimental evidence proves that frameshifting is responsible for the *bldA* phenotype, experiments should then address the premise that TTA codons modulate gene expression in wild-type systems. As these experiments will require highly sensitive detection methods as frameshifting in *bldA*+ strains is predicted to occur at low levels, preliminary experimentation is required to determine the limits of detection of frameshift and in-frame mistranslation products.

Currently, several of the proposed experiments are not possible since the $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{CAA}}$ and $\mathrm{tRNA}^{\mathrm{Leu}}_{\mathrm{UAG}}$ as well as many other tRNA genes such as, $\mathrm{tRNA}^{\mathrm{Tyr}}_{\mathrm{GUA}}$ have not yet been identified in *Streptomyces*. In addition to the requirement for tRNA gene sequences, purified preparations of individual tRNA species would also be useful for additional *in* *vivo* and *in vitro* experiments to determine what additional sequences are "shifty", what tRNAs carry out these shifts, and under what conditions these shifts occur. The identification of modified bases in *Streptomyces* tRNAs is also important for predicting mistranslation, and has implications for the ability to cause a frameshift. So while this work has answered a few questions regarding the role of *bldA* in *Streptomyces*, more importantly, it has generated more questions that remain to be addressed.

V. BIBLIOGRAPHY

Aceti, D. J., W. C. Champness (1998) Transcriptional regulation of *Streptomyces coelicolor* pathway-specific antibiotic regulators by the *absA* and *absB* loci. J. Bacteriol. 180: 3100-3106.

Adamidis, T., W. Champness (1992) Genetic analysis of *absB*, a *Streptomyces coelicolor* locus involved in global antibiotic regulation. J. Bacteriol. 174: 4622-4628.

Adamidis, T., P. Riggle, W. Champness (1990) Mutations in a new *Streptomyces coelicolor* locus which globally block antibiotic biosynthesis but not sporulation. J. Bacteriol. 172: 2962-2969.

Aharonowitz, Y., G. Cohen, J. F. Martin (1992) Penicillin and cephalosporin biosynthetic genes: structure, organization, regulation, and evolution. Annu. Rev. Microbiol. 46: 461-495.

Aidoo, K. A., A. Wong, D. C. Alexander, R. A. R. Rittammer, S. E. Jensen (1994) Cloning, sequencing and disruption of a gene from *Streptomyces clavuligerus* involved in clavulanic acid biosynthesis. Gene 147: 41-46.

Alexander, D. C. (1998) Regulation of cephamycin biosynthesis in *Streptomyces clavuligerus*. Ph.D. Thesis, Department of Biological Sciences, University of Alberta, Edmonton, Alberta, Canada

Alexander, D. C., S. E. Jensen (1998) Investigation of the *Streptomyces clavuligerus* cephamycin C gene cluster and its regulation by the CcaR protein. J. Bacteriol. 180: 4068-4079.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers, D. J. Lipman (1990) Basic local alignment search tool. J. Mol. Biol. 215: 403-410.

Amersham Manual (1985) Membrane transfer and detection methods. Oakville, Ont.: Amersham International.

Arcamone, F., G. Cassinelli, G. Fantini, A. Grein, P. Orezzi, C. Pol, C. Spalla (1969) Adriamycin, 14-hydroxydaunomycin, a new antitumor antibiotic from *S. peucetius* var. *caesius*. Biotech. Bioeng. 11: 1101-1110.

Arisawa, A., N. Kawamura, H. Tsunekawa, K. Okamura, H. Tone, R. Okamoto (1993) Cloning and nucleotide sequences of two genes involved in the 4"-*O*-acylation of macrolide antibiotics from *Streptomyces thermotolerans*. Biosci. Biotech. Biochem. 57: 2020-2025.

Arisawa, A., H. Tsunekawa, K. Okamura, R. Okamoto (1995) Nucleotide sequence analysis of the carbomycin biosynthetic genes including the 3'-*O*-acyltransferase gene from *Streptomyces thermotolerans*. Biosci. Biotech. Biochem. 59: 582-588.

Atkins, J. F., R. F. Gesteland (1995) Discontinuous triplet decoding with or without re-pairing by peptidyl tRNA. *In* tRNA: Structure, Biosynthesis, and Function. Edited by D. Soll and U. RajBhandary. Washington, DC: American Society for Microbiology, pp. 471-490.

Atkins, J. F., R. F. Gesteland, B. R. Reid, C. W. Anderson (1979) Normal tRNAs promote ribosomal frameshifting. Cell 18: 1119-1131.

August, P. R., M. C. Flickinger, D. H. Sherman (1994) Cloning and analysis of a locus (mcr) involved in mitomycin C resistance in *Streptomyces*. J. Bacteriol. 176: 4448.

Babcock, M. J., K. E. Kendrick (1990) Transcriptional and translational features of a sporulation gene of *Streptomyces griseus*. Gene 95: 57-63.

Bachmann, B. O., R. Li, C. A. Townsend (1998) β-lactam synthetase: a new biosynthetic enzyme. Proc. Natl. Acad. Sci. USA 95: 9082-9086.

Bailey, C. R., D. J. Winstanley (1986) Inhibition of restriction in *Streptomyces clavuligerus* by heat treatment. J. Gen. Microbiol. 132: 2945-2947.

Bar-Nir, D., A. Cohen, M. E. Goedeke (1992) tDNA^{Ser} sequences are involved in the excision of *Streptomyces griseus* plasmid pSG1. Gene 122: 71-76.

Beall, B., J. Lutkenhaus (1991) FtsZ in Bacillus subtilis is required for vegetative septation and for asymmetric septation during sporulation. Genes Dev. 5: 447-455.

Belcourt, M. F., P. J. Farabaugh (1990) Ribosomal frameshifting in the yeast retrotransposon Ty: tRNAs induce slippage on a 7 nucleotide minimal site. Cell 62: 339-352.

Beyer, S., J. Distler, W. Piepersberg (1996) The *str* gene cluster for the biosynthesis of 5'-hydroxystreptomycin in *Streptomyces glaucescens* GLA.0 (ETH 22794): new operons and evidence for pathway-specific regulation by StrR. Mol. Gen. Genet. 250: 775-784.

Bibb, M. (1996) The regulation of antibiotic production in *Streptomyces coelicolor* A3(2). Microbiology 142: 1335-1344.

Bibb, M. J., P. R. Findlay, M. W. Johnson (1984) The relationship between base composition and codon usage in bacterial genes and its use in the simple and reliable identification of protein coding sequences. Gene 30: 157-166.

Bibb, M. J., G. R. Janssen, J. M. Ward (1985) Cloning and analysis of the promoter region of the erythromycin resistance gene *(ermE)* of *Streptomyces erythraeus*. Gene 38: 215-226.

Bibb, M. J., G. R. Janssen, J. M. Ward (1986) Cloning and analysis of the promoter region of the erythromycin resistance gene *(ermE)* of *Streptomyces erythraeus*. Gene 41: E357-E368.

Bierman, M., R. Logan, K. O'Brien, E. T. Seno, N. Rao, B. E. Schoner (1992) Plasmid cloning vectors for the conjugal transfer of DNA from *Escherichia coli* to *Streptomyces* spp. Gene 116: 43-49.

Biro, S., I. Bekesi, S. Vitalis, G. Szabo (1980) A substance effecting differentiation in *Streptomyces griseus*. Eur. J. Biochem. 103: 359-363.

Bjork, G. R. (1995) Biosynthesis and Function of Modified Nucleosides. *In* tRNA: Structure, Biosynthesis, and Function. Edited by D. Soll and U. RajBhandary. Washington, DC: American Society for Microbiology, pp. 165-205.

Bourn, W. R., B. Babb (1995) Computer assisted identification and classification of streptomycete promoters. Nucl. Acids Res. 23: 3696-3703.

Bradford, M. M. (1976) A rapid and sensitive method for the quantification of microgram quantities of proteins utilizing the principle of protein dye-binding. Anal. Biochem. 72: 248-254.

Brian, P., P. J. Riggle, R. A. Santos, W. C. Champness (1996) Global negative regulation of *Streptomyces coelicolor* antibiotic synthesis mediated by an *absA*-encoded putative signal transduction system. J. Bacteriol. 178: 3221-3231.

Brierley, I., A. J. Jenner, S. C. Inglis (1992) Mutational analysis of the "slippery-sequence" component of a coronavirus ribosomal frameshifting signal. J. Mol. Biol. 227: 463-479.

Brown, D., J. R. Evans (1979) Structures of three novel β-lactams isolated from *Streptomyces clavuligerus*. J.C.S. Chem. Comm. 1979: 282-283.

Brown, M. P., K. A. Aidoo, L. C. Vining (1996) A role for *pabAB*, a *p*-aminobenzoate synthase gene of *Streptomyces venezuelae* ISP5230, in chloramphenicol biosynthesis. Microbiology 142: 1345-1355.

Bullock, W. O., J. M. Fernandez, J. M. Short (1987) XL1-Blue: A high efficiency plasmid transforming *recA Escherichia coli* strain with beta-galactosidase selection. BioTechniques 5: 376-378.

Bulmer, M. (1987) Coevolution of codon usage and transfer RNA abundance. Nature 325: 728-730.

Caballero, J. L., F. Malpartida, D. A. Hopwood (1991) Transcriptional organization and regulation of an antibiotic export complex in the producing *Streptomyces* culture. Mol. Gen. Genet. 228: 372-380.

Chaconas, G., J. H. van de Sande (1980) 5'-^{32}P labeling of RNA and DNA restriction fragments. Methods Enzymol. 65: 75-85.

Champness, W., P. Riggle, T. Adamidis, P. Vandervere (1992) Identification of *Streptomyces coelicolor* genes involved in regulation of antibiotic synthesis. Gene 115: 55-60.

Champness, W. C. (1988) New loci required for *Streptomyces coelicolor* morphological and physiological differentiation. J. Bacteriol. 170: 1168-1174.

Chang, H.-M., M.-Y. Chen, Y.-T. Shieh, M. J. Bibb, C. W. Chen (1996) The *cutRS* signal transduction system of *Streptomyces lividans* represses the biosynthesis of the polyketide antibiotic actinorhodin. Mol. Microbiol. 21: 1075-1085.

Chater, K. F. (1972) A morphological and genetic mapping study of white colony mutants of *Streptomyces coelicolor*. J. Gen. Microbiol. 72: 9-28.

Chater, K. F. (1975) Construction and phenotypes of double sporulation deficient mutants in *Streptomyces coelicolor* A3(2). J. Gen. Microbiol. 87: 312-325.

Chater, K. F. (1989) Sporulation in *Streptomyces*. *In* Regulation of Prokaryotic Development. Edited by I. Smith, R. A. Slepecky and P. Setlow. Washington, D.C.: American Society for Microbiology, pp. 277-299.

Chater, K. F. (1993) Genetics of differentiation in *Streptomyces*. Annu. Rev. Microbiol. 47: 685-713.

Chater, K. F., C. J. Bruton, A. A. King, J. E. Suárez (1982) The expression of *Streptomyces* and *Escherichia coli* drug resistance determinants cloned into the *Streptomyces* phage ϕC31. Gene 19: 21-32.

Chater, K. F., C. J. Bruton, K. A. Plaskitt, M. J. Buttner, C. Mendez, J. D. Helmann (1989) The developmental fate of *S. coelicolor* hyphae depends upon a gene product homologous with the motility σ factor of *B. subtilis*. Cell 59: 133-143.

Chater, K. F., M. J. Merrick (1979) Streptomycetes. *In* Developmental Biology of Prokaryotes. Edited by J. H. Parish. Oxford: Blackwell Scientific Publications, Ltd., pp. 93-114.

Chen, G.-F., M. Inouye (1990) Suppression of the negative effect of minor arginine codons on gene expression; preferential usage of minor codons within the first 25 codons of the *Escherichia coli* genes. Nucl. Acids Res. 18: 1465-1473.

Chen, G.-F. T., M. Inouye (1994) Role of the AGA/AGG codons, the rarest codons in global gene expression in *Escherichia coli*. Genes Dev. 8: 2641-2652.

Chong, P. P., S. M. Podmore, H. M. Kieser, M. Redenbach, K. Turgay, M. Marahiel, D. A. Hopwood, C. P. Smith (1998) Physical identification of a chromosomal locus encoding biosynthetic genes for the lipopeptide calcium-dependent antibiotic (CDA) of *Streptomyces coelicolor* A3(2). Microbiology 144: 193-199.

Clayton, T. M., M. J. Bibb (1990) *Streptomyces* promoter-probe plasmids that utilise the *xylE* gene of *Pseudomonas putida*. Nucl. Acids Res. 18: 1077.

Coque, J. J. R., J. F. Martin, J. G. Calzada, P. Liras (1991) The cephamycin biosynthetic genes *pcbAB*, encoding a large multidomain peptide synthetase, and *pcbC* of *Nocardia lactamdurans* are clustered together in an organization different from the same genes in *Acremonium chrysogenum* and *Penicillium chrysogenum*. Mol. Microbiol. 5: 1125-1133.

Cormack, B. P., R. H. Valdivia, and, S. Falkow (1996) FACS-optimized mutants of the green fluorescent protein (GFP). Gene 173: 33-38.

Crandall, L. W., R. L. Hamill (1986) Antibiotics produced by *Streptomyces:* major structural classes. *In* Antibiotic-producing *Streptomyces*, Volume 9. Edited by I. C. Gunsalus and R. Y. Stanier. New York: Academic Press, Inc., pp. 355-401.

Crick, F. H. C. (1966) Codon-anticodon pairing: the wobble hypothesis. J. Mol. Biol. 19: 548-555.

Curran, J. F. (1995) Decoding with the A:I wobble pair is inefficient. Nucl. Acids Res. 23: 683-688.

Dai, K., J. Lutkenhaus (1991) *ftsZ* is an essential cell division gene in *Escherichia coli*. J. Bacteriol. 173: 3500-3506.

Davis, N. K., K. F. Chater (1990) Spore colour in *Streptomyces coelicolor* A3(2) involves the developmentally regulated synthesis of a compound biosynthetically related to polyketide antibiotics. Mol. Microbiol. 4: 1679-1691.

Davis, N. K., K. F. Chater (1992) The *Streptomyces coelicolor whiB* gene encodes a small transcription factor-like protein dispensable for growth but essential for sporulation. Mol. Gen. Genet. 232: 351-358.

Dayhuff, T. J., J. F. Atkins, R. F. Gesteland (1986) Characterization of ribosomal frameshift events by protein sequences analysis. J. Biol. Chem. 261: 7491-7500.

de la Fuente, J. L., A. Rumbero, J. F. Martin, P. Liras (1997) Δ-1-piperideine-6-carboxylate dehydrogenase, a new enzyme that forms α-aminoadipate in *Streptomyces clavuligerus* and other cephamycin C-producing actinomycetes. Biochem J. 327: 59-64.

Decker, H., C. R. Hutchinson (1993) Transcriptional analysis of the *Streptomyces glaucescens* tetracenomycin C biosynthesis gene cluster. J. Bacteriol. 175: 3887-3892.

Decker, H., H. Motamedi, C. R. Hutchinson (1993) Nucleotide sequences and heterologous expression of *tcmG* and *tcmP*, biosynthetic genes for tetracenomycin C in *Streptomyces glaucescens*. J. Bacteriol. 175: 3876-3886.

DeHoff, B. S., K. L. Sutton, P. R. J. Rosteck (1997) Unpublished.

Demidchuk, J. A., V. A. Fedorenko, Z. V. Oliynyk (1997) Unpublished.

Dharmatilake, A. J., K. E. Kendrick (1994) Expression of the division-controlling gene *ftsZ* during growth and sporulation of the filamentous bacterium *Streptomyces griseus*. Gene 147: 21-28.

Di Marco, A., M. Gaetani, P. Orezzi, B. M. Scarpinato, R. Silvestrini, M. Soldati, T. Dasdia, L. Valentini (1964) 'Daunomycin', a new antibiotic of the rhodomycin group. Nature 201: 706-707.

Dickens, M. L., W. R. Strohl (1996) Isolation and characterization of a gene from *Streptomyces* sp. Strain C5 that confers the ability to convert daunomycin to doxorubicin on *Streptomyces lividans* TK24. J. Bacteriol. 178: 3389-3395.

Dirheimer, G., G. Keith, P. Dumas, E. Westhof (1995) Primary, Secondary, and Tertiary Structures of tRNAs. *In* tRNA: Structure, Biosynthesis, and Function. Edited by D. Soll and U. RajBhandary. Washington, DC: American Society for Microbiology, pp. 93-126.

Distler, J., A. Ebert, K. Mansouri, K. Pissowotzki, M. Stockmann, W. Piepersberg (1987) Gene cluster for streptomycin biosynthesis in *Streptomyces griseus*: Nucleotide sequence of three genes and analysis of transcriptional activity. Nucleic Acids Res. 15: 8041-8056.

Distler, J., K. Mansouri, G. Mayer, M. Stockmann, W. Piepersberg (1992) Streptomycin biosynthesis and its regulation in Streptomycetes. Gene 115: 105-111.

Doran, J. L., B. K. Leskiw, S. Aippersbach, S. E. Jensen (1990) Isolation and characterization of a β-lactamase-inhibitory protein from *Streptomyces clavuligerus* and cloning and analysis of the corresponding gene. J. Bacteriol. 172: 4909-4918.

Durairaj, M., B. K. Leskiw, S. E. Jensen (1996) Genetic and biochemical analysis of the cysteinyl residues of isopenicillin N synthase from *Streptomyces clavuligerus*. Can. J. Microbiol. 42: 870-875.

Egan, L. A., R. W. Busby, D. Iwata-Reuyl, C. A. Townsend (1997) Probable role of clavaminic acid as the terminal intermediate in the common pathway to clavulanic acid and the antipodal clavam metabolites. J. Am. Chem. Soc. 119: 2348-2355.

Elliot, M., F. Damji, R. Passantino, K. Chater, B. Leskiw (1998) The *bldD* gene of *Streptomyces coelicolor* A3(2): a regulatory gene involved in morphogenesis and antibiotic production. J. Bacteriol. 180: 1549-1555.

Emilsson, V., C. G. Kurland (1990) Growth rate dependence of transfer RNA abundance in *Escherichia coli*. EMBO J. 9: 4359-4366.

Epp, J. K., S. G. Burgett, B. E. Schoner (1987) Cloning and nucleotide sequence of a carbomycin-resistance gene from *Streptomyces thermotolerans*. Gene 53: 73-83.

Farabaugh, P. J., G. R. Bjork (1999) How translation accuracy influences reading frame maintenance. EMBO J. 18: 1427-1434.

Farabaugh, P. J., H. Zhao, A. Vimaladithan (1993) A novel programed frameshift expresses the *POL3* gene of retrotransposon Ty3 of yeast: frameshifting without tRNA slippage. Cell 74: 93-103.

Feinberg, A. P., B. Vogelstein (1983) A technique for radiolabeling DNA restriction endonuclease fragments to high specific activity. Anal. Biochemistry 132: 6-13.

Fernández-Moreno, M. A., J. L. Caballero, D. A. Hopwood, F. Malpartida (1991) The *act* cluster contains regulatory and antibiotic export genes, direct targets for translational control by the *bldA* transfer RNA gene of *Streptomyces*. Cell 66: 769-780.

Fernández-Moreno, M. A., A. J. Martin-Triana, E. Martinez, J. Niemi, H. M. Kieser, D. A. Hopwood, F. Malpartida (1992) *abaA*, a new pleiotropic regulatory locus for antiobiotic production in *Streptomyces coelicolor*. J. Bacteriol. 174: 2958-2967.

Ferrari, F. A., K. Trach, D. LeCoq, J. Spence, E. Ferrari, J. A. Hoch (1985) Characterization of the *spo0A* locus and its deduced product. Proc. Natl. Acad. Sci. USA 82: 2647-2651.

Folley, L. S., M. Yarus (1989) Codon contexts from weakly expressed genes reduce expression *in vivo*. J. Mol. Biol. 209: 359-378.

Foulstone, M., C. Reading (1982) Assay of amoxicillin and clavulanic acid, the components of augmentin, in biological fluids with high-performance liquid chromatography. Antimicrob. Agents Chemother. 22: 753-762.

Fournier, M. J., H. Ozeki (1985) Structure and organization of the transfer ribonucleic acid genes of *Escherichia coli* K-12. Microbiol. Rev. 49: 379-397.

Frost, L. S., S. Lee, N. Yanchar, W. Paranchych (1989) Antisense RNA suggest a model for FinOP action in the repression of bacterial conjugation by the F*lac* plasmid JCFLO. Mol. Gen. Genet. 218: 152-160.

Gandecha, A. R., E. Cundliffe (1996) Molecular analysis of *tlrD*, an MLS resistance determinant from the tylosin producer, *Streptomyces fradiae*. Gene 180: 173-176.

Gandecha, A. R., S. L. Large, E. Cundliffe (1997) Analysis of four tylosin biosynthetic genes from the *tylLM* region of the *Streptomyces fradiae* genome. Gene 184: 197-203.

Garcia, G. M., P. K. Mar, D. A. Mullin, J. R. Walker, N. E. Prather (1986) The *E. coli dnaY* gene encodes an arginine transfer RNA. Cell 45: 453-459.

Garda, A. L., J. M. Fernandex-Abalos, P. Sanchez, A. Ruiz-Arribas, R. I. Santamaria (1997) Two genes encoding an endoglucanase and a cellulose-binding protein are clustered and co-regulated by a TTA codon in *Streptomyces halstedii* JM8. Biochem. J. 324: 402-411.

Geistlich, M., R. Losick, J. R. Turner, R. N. Rao (1992) Characterization of a novel regulatory gene governing the expression of a polyketide synthase gene in *Streptomyces ambofaciens*. Mol. Microbiol. 6: 2019-2029.

Giebelhaus, L. A., L. Frost, E. Lanka, E. P. Gormley, J. E. Davies, B. Leskiw (1996) The Tra2 core of the IncPα plasmid RP4 is required for intergeneric mating between *Escherichia coli* and *Streptomyces lividans*. J. Bacteriol. 178: 6378-6381.

Grafe, U., G. Reinhardt, W. Schade, I. Eritt, W. F. Fleck, L. Radics (1983) Interspecific inducers of cytodifferentiation and anthracycline biosynthesis from *Streptomyces bikinensis* and *S. cyaneofuscatus*. Biotech. Lett. 5: 591-596.

Graffe, M., J. Dondon, J. Caillet, P. Romby, C. Ehresmann, B. Ehresmann, M. Springer (1992) The specificity of translation control switched with transfer RNA identity rules. Science 255: 994-996.

Gramajo, H. C., E. Takano, M. J. Bibb (1993) Stationary phase production of the antibiotic actinorhodin in *Streptomyces coelicolor* A3(2) is transcriptionally regulated. Mol. Microbiol. 7: 837-845.

Guijarro, J., R. Santamaria, A. Schauer, R. Losick (1988) Promoter determining the timing and spatial localization of transcription of a cloned *Streptomyces coelicolor* gene encoding a spore-associated polypeptide. J. Bacteriol. 170: 1895-1901.

Guthrie, E. P., K. F. Chater (1990) The level of a transcript required for the production of a *Streptomyces coelicolor* antibiotic is conditionally dependent on a transfer RNA gene. J. Bacteriol. 172: 6189-6193.

Guthrie, E. P., C. S. Flaxman, J. White, D. A. Hodgson, M. J. Bibb, K. F. Chater (1998) A response-regulator-like activator of antibiotic synthesis from *Streptomyces coelicolor* A3(2) with an amino-terminal domain that lacks a phosphorylation pocket. Microbiology 144: 727-738.

Hagege, J., F. Boccard, T. Smokvina, J.-L. Pernodet, A. Friedmann, M. Guerineau (1994) Identification of a gene encoding the replication initiator protein of the *Streptomyces* integrating element, pSAM2. Plasmid 31: 166-183.

Hagege, J., J.-L. Pernodet, A. Friedmann, M. Guerineau (1993) Mode and origin of replication of pSAM2, a conjugative integrating element of *Streptomyces ambofaciens*. Mol. Microbiol. 10: 799-812.

Hara, O., T. Beppu (1982) Mutants blocked in streptomycin production in *Streptomyces griseus* - the role of A-factor. J. Antibiot. 35: 349-358.

Hara, O., S. Horinouchi, T. Uozumi, T. Beppu (1983) Genetic analysis of A-factor synthesis in *Streptomyces coelicolor* A3(2) and *Streptomyces griseus*. J. Gen. Microbiol. 129: 2939-2914.

Hara, O., C. R. Hutchinson (1990) Cloning of midecamycin (MLS)-resistance genes from *Streptomyces mycarofaciens, Streptomyces lividans* and *Streptomyces coelicolor* A3(2). J. Antibiot. 63: 977-991.

Hara, O., C. R. Hutchinson (1992) A Macrolide 3-*O*-acyltransferase gene from the midecamycin-producing species *Streptomyces mycarofaciens*. J. Bacteriol. 174: 5141-5144.

Hashimoto, K., T. Nihira, S. Sakuda, Y. Yamada (1992) IM-2, a butyrolactone autoregulator, induces production of several nucleoside antibiotics in *Streptomyces* sp. FRI-5. J. Ferment. Bioeng. 73: 449-455.

Heinzel, P., O. Werbitzky, J. Distler, W. Piepersberg (1988) A second streptomycin resistance gene from *Streptomyces griseus* codes for streptomycin-3"-phosphotransferase. Arch. Microbiol. 150: 184-192.

Herron, P. R., M. C. Evans, P. J. Dyson (1998) Insertion site specificity of IS6100 in *Streptomyces*. Unpublished.

Hillemann, D., A. Pühler, W. Wohlleben (1991) Gene disruption and gene replacement in *Streptomyces* via single stranded DNA transformation of integration vectors. Nucl. Acids Res. 19: 727-731.

Hirakata, T., H. Kieser, D. Hopwood, H. Urabe, H. Ogarawa (1998) Putative protein serine/threonine kinase gene are located in several positions on the chromosome of *Streptomyces coelicolor* A3(2). FEMS Microbiol. Lett. 159: 1-5.

Hobbs, G., A. I. C. Obanye, J. Petty, C. Mason, E. Barratt, D. C. J. Gardner, F. Flett, C. P. Smith, P. Broda, S. G. Oliver (1992) An intergrated approach to studying regulation of production of the antibiotic methylenomycin by *Streptomyces coelicolor* A3(2). J. Bacteriol. 174: 1487-1494.

Hoch, J. A. (1993) Regulation of the phosphorelay and the initiation of sporulation in *Bacillus subtilis*. Annu. Rev. Microbiol. 47: 441-465.

Hodgson, J. E., A. P. Fosberry, N. S. Rawlinson, H. N. M. Ross, R. J. Neal, J. C. Arnell, A. J. Earl, E. J. Lawlor (1995) Clavulanic acid biosynthesis in *Streptomyces clavuligerus:* gene cloning and characterization. Gene 166: 49-55.

Homerova, D., Z. Hollanderova, J. Kormanec, J. Sevcik (1992) Cloning and sequencing of the gene encoding a ribonuclease from *Streptomyces aureofaciens* CCM3239. Gene 119: 147-148.

Hong, S.-K., A. Matsumoto, S. Horinouchi, T. Beppu (1993) Effects of protein kinase inhibitors on in vitro protein phosphorylation and cellular differentiation of *Streptomyces griseus*. Mol. Gen. Genet. 236: 347-354.

Hopwood, D. A. (1978) Extrachromosomally determined antibiotic production. Annu. Rev. Microbiol. 32: 373-392.

Hopwood, D. A., M. J. Bibb, K. F. Chater, T. Kieser, C. J. Bruton, H. M. Kieser, D. J. Lydiate, C. P. Smith, J. M. Ward, H. Schrempf (1985) Genetic manipulation of *Streptomyces:* A Laboratory Manual.: John Innes Foundation, Norwich, UK.

Hopwood, D. A., T. Kieser, H. M. Wright, M. J. Bibb (1983) Plasmids, recombination and chromosomal mapping in *Streptomyces lividans* 66. J. Gen. Microbiol. 129: 2257-2269.

Hopwood, D. A., D. H. Sherman (1990) Molecular genetics of polyketides and its comparison to fatty acid biosynthesis. Annu. Rev. Genet. 24: 37-66.

Hopwood, D. A., H. Wildermuth, H. M. Palmer (1970) Mutants of *Streptomyces coelicolor* defective in sporulation. J. Gen. Microbiol. 61: 397-408.

Hopwood, D. A., H. M. Wright (1983) CDA is a new chromosomally-determined antibiotic from *Streptomyces coelicolor* A3(2). J. Gen. Microbiol. 129: 2575-3579.

Horii, M., T. Ishizaki, S.-Y. Paik, T. Manome, Y. Murooka (1990) An operon containing the genes for cholesterol oxidase and a cytochrome P-450-like protein from a *Streptomyces* sp. J. Bacteriol. 172: 3644-3653.

Horinouchi, S. (1986) Nucleotide sequence of *afsB*, a pleiotropic gene involved in secondary metabolism in *Streptomyces coelicolor* A3(2) and *Streptomyces lividans*. J. Bacteriol. 168: 257-269.

Horinouchi, S. (1993) "Eukaryotic" signal transduction systems in the bacterial genus *Streptomyces*. Actinomycetol 7: 68-87.

Horinouchi, S., T. Beppu (1990a) Autoregulatory factors of secondary metabolism and morphogenesis in actinomycetes. Biotechnology 10: 191-204.

Horinouchi, S., T. Beppu (1992b) Autoregulatory factors and communication in actinomycetes. Annu. Rev. Microbiol. 46: 377-398.

Horinouchi, S., O. Hara, T. Beppu (1983) Cloning of a pleiotropic gene that positively controls biosynthesis of A-factor, actinorhodin, and prodigiosin in *Streptomyces coelicolor* A3(2) and *Streptomyces lividans*. J. Bacteriol. 155: 1238-1248.

Hunter, P. A., K. Coleman, J. Fisher, D. Taylor (1980) *In vitro* synergistic properties of clavulanic acid, with ampicillin, amoxycillin and ticarcillin.
J. Antimicrob. Chemother. 6: 455-470.

Ikemura, T. (1981a) Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the respective codons in its protein genes.
J. Mol. Biol. 146: 1-21.

Ikemura, T. (1981b) Correlation between the abundance of *Escherichia coli* transfer RNAs and the occurrence of the repective codons in its protein genes: a proposal for a synonymous codon choice that is optimal for the *E. coli* translational system.
J. Mol. Biol. 151: 389-409.

Inagaki, Y., A. Kojima, Y. Bessho, H. Hori, T. Ohama, S. Osawa (1995) Translation of synonymous codons in family boxes by *Mycoplasma capricolum* tRNAs with unmodified uridine or adenosine at the first anticodon position. J. Mol. Biol. 251: 486-492.

Ingram, C., M. Brawner, P. Youngman, J. Westpheling (1989) *xylE* functions as an efficient reporter gene in *Streptomyces* spp.: use for the study of *galP1*, a catabolite-controlled promoter. J. Bacteriol. 171: 6617-6624.

Ishizuka, H., S. Horinouchi, H. M. Kieser, D. A. Hopwood, T. Beppu (1992) A putative two-component regulatory system involved in secondary metabolism in *Streptomyces* spp. J. Bacteriol. 174: 7585-7594.

Itoh, M., A. Penyige, S. Okamoto, K. Ochi (1996) Proteins that interact with GTP in *Streptomyces griseus* and its possible implication in morphogenesis.
FEMS Microbiol. Lett. 135: 311-316.

Jacks, T., H. D. Madhanik, F. R. Masiarz, H. E. Varmus (1988a) Signals for ribosomal frameshifting in the Rous Sarcoma virus *gag-pol* region. Cell 55: 447-458.

Jacks, T., M. D. Power, F. R. Masiarz, P. A. Luciw, P. J. Barr, H. E. Varmus (1988b) Characterisation of ribosomal frameshifting in HIV-1 *gag-pol* expression.
Nature 331: 280-283.

Janc, J. W., L. A. Egan, C. A. Townsend (1993) Emerging evidence for a shared biosynthetic pathway among clavulanic acid and the structurally diverse clavam metabolites. Bioorg. Med. Chem. Lett. 3: 2313-2316.

Jensen, S. E., A. L. Demain (1993) β-lactams. *In* Biochemistry and Genetics of Antibiotic Biosynthesis. Edited by L. C. Vining. Stoneham, MA.: Butterwoth Heinemann, pp. 71-99.

Kakavas, S. J., L. Katz, D. Stassi (1997) Identification and characterization of the niddamycin polyketide synthase genes from *Streptomyces caelestis*. J. Bacteriol. 179: 7515-7522.

Katayama, M., Y. Sakai, S. Okamoto, F. Ihara, T. Nihira, Y. Yamada (1996) Gene organization in the *ada-rplL* region of *Streptomyces virginiae*. Gene 171: 135-136.

Kawamoto, S., M. Nakamura, S. Yashima (1993) Cloning, sequence and expression of the tyrosinase gene from *Streptomyces lavendulae* MA406 A-1. J. Ferment. Bioeng. 76: 345-355.

Kawamoto, S., H. Watanabe, A. Hesketh, J. C. Ensign, K. Ochi (1997) Expression analysis of the *ssgA* gene product, associated with sporulation and cell division in *Streptomyces griseus*. Microbiology 143: 1077-1086.

Kelemen, G. H., P. Brian, K. Flardh, L. Chamberlin, K. F. Chater, M. J. Buttner (1998) Developmental regulation of transcription of *whiE*, a locus specifying the polyketide spore pigment in *Streptomyces coelicolor* A3(2). J. Bacteriol. 180: 2515-2521.

Kelemen, G. H., G. L. Brown, J. Kormanec, L. Potuckova, K. F. Chater, M. J. Buttner (1996) The positions of the sigma-factor genes, *whiG* and *sigF* in the hierarchy controlling the development of spore chains in the aerial hyphae of *Streptomyces coelicolor* A3(2). Mol. Microbiol. 21: 593-603.

Kelemen, G. H., K. A. Plaskitt, C. G. Lewis, K. C. Findlay, M. J. Buttner (1995) Deletion of DNA lying close to the *glkA* locus induces ectopic sporulation in *Streptomyces coelicolor* A3(2). Mol. Microbiol. 17: 221-230.

Kenig, M., C. Reading (1979) Holomycin and an antibiotic (MM19290) related to tunicamycin, metabolites of *Streptomyces clavuligerus*. J. Antibiot. 32: 549-554.

Kim, S., K. L. Lee (1996) Trypsin-like protease of *Streptomyces exfoliatus* SMF13, a potential agent in mycelial differentiation. Microbiology 142: 1797-1806.

Kirby, K. S., E. Fox-Carter, M. Guest (1967) Isolation of deoxyribonucleic acid and ribosomal ribonucleic acid from bacteria. Biochem. J. 104: 258-262.

Kitchingman, G. R., M. J. Fournier (1975) Unbalanced growth and the production of unique transfer ribonucleic acids in relaxed-control *Escherichia coli*. J. Bacteriol. 124: 1382-1394.

Knowles, J. R. (1985) Penicillin resistance: the chemistry of β-lactamase inhibition. Acc. Chem. Res. 18: 97-104.

Kondo, S., K. Yasui, M. Natsume, M. Katayama, S. Marumo (1988) Isolation, physico-chemical properties and biological activity of pamamycin-607, an aerial mycelium-inducing substance form *Streptomyces alboniger*. J. Antibiot. 41: 1196-1204.

Korn-Wendisch, F., H. J. Kutzner (1992) The Family Streptomycetaceae. *In* The Prokaryotes: a handbook on the biology of bacteria: ecophysiology, isolation, identification, applications. Edited by A. Balows. New York: Springer-Verlag, pp. 921-980.

Kovacevic, S., M. B. Tobin, J. R. Miller (1990) The β-lactam biosynthesis genes for isopenicillin N epimerase and deacetoxycephalosporin C synthetase are expressed from a single transcript in *Streptomyces clavuligerus*. J. Bacteriol. 172: 3952-3958.

Kudo, N., M. Kimura, T. Beppu, S. Horinouchi (1995) Cloning and characterization of a gene involved in aerial mycelium formation in *Streptomyces griseus*. J. Bacteriol. 177: 6401-6410.

Kurland, C. G., M. Ehrenberg (1984) Optimization of translation accuracy. Prog. Nucleic Acid Res. Mol. Biol. 31: 191-219.

Kwak, J., L. A. McCue, K. E. Kendrick (1996) Identification of *bldA* mutants of *Streptomyces griseus*. Gene 171: 75-78.

Lakey, J. H., E. J. A. Lea, B. A. Rudd, H. M. Wright, D. A. Hopwood (1983) A new channel-forming antibiotic from *Streptomyces coelicolor* A3(2) which requries calcium for its activity. J. Gen. Microbiol. 129: 3565-3573.

Lawlor, E. J. (1987) Molecular Genetics of *bldA*, a developmental gene of *Streptomyces coelicolor*: Ph. D. Thesis, University of East Anglia, Norwich, England.

Lawlor, E. J., H. A. Baylis, K. F. Chater (1987) Pleiotropic morphological and antibiotic deficiencies result from mutations in a gene encoding a tRNA-like product in *Streptomyces coelicolor* A3(2). Genes Dev. 1: 1305-1310.

Leclerc, G., C. Sirard, G. R. Drapeau (1989) the *Escherichia coli* cell division mutation *ftsM1* is in *serU*. J. Bacteriol. 171: 2090-2095.

Leskiw, B. K., M. J. Bibb, K. F. Chater (1991a) The use of a rare codon specifically during development. Mol. Microbiol. 5: 2861-2867.

Leskiw, B. K., E. J. Lawlor, J. M. Fernandez-Abalos, K. F. Chater (1991b) TTA codons in some genes prevent their expression in a class of developmental, antibiotic-negative, *Streptomyces* mutants. Proc. Natl. Acad. Sci. USA 88: 2461-2465.

Leskiw, B. K., R. Mah (1995) The *bldA*-encoded tRNA is poorly expressed in the *bldI* mutant of *Streptomyces coelicolor* A3(2). Microbiology 141: 1921-1926.

Leskiw, B. K., R. Mah, E. J. Lawlor, K. F. Chater (1993) Accumulation of *bldA*-specified tRNA is temporally regulated in *Streptomyces coelicolor* A3(2). J. Bacteriol. 175: 1995-2005.

Li, Y., W. R. Strohl (1996) Cloning, purification, and properties of a phosphotyrosine protein phosphatase from *Streptomyces coelicolor* A3(2). J. Bacteriol. 178: 136-142.

Lim, V. I. (1995) Analysis of action of the wobble adenine on codon reading within the ribosome. J. Mol. Biol. 252: 277-282.

Lindsley, D., J. Gallant (1993) On the directional specificity of ribosome frameshifting at a "hungry" codon. Proc. Natl. Acad. Sci. USA 90: 5469-5473.

Lombo, F., A. F. Brana, C. Mendez, J. A. Salas (1999) The mithramycin gene cluster of *Streptomyces argillaceus* contains a positive regulatory gene and two repeated DNA sequences that are located a both ends of the cluster. J. Bacteriol. 181: 642-647.

Lomovskaya, N. D., K. F. Chather, N. M. Mkrtumian (1980) Genetics and molecular biology of *Streptomyces* bacteriophages. Microbiol. Rev. 44: 206-229.

Looman, A. C., J. Bodlaender, L. J. Comstock, D. Eaton, P. Jhurani, H. A. de Boer, P. H. van Knippenberg (1987) Influence of the codon following the AUG initiation codon on the expression of a modified *lacZ* gene in *Escherichia coli*. EMBO J. 6: 2489-2492.

Losick, R., J. Pero (1981) Cascades of sigma factors. Cell 25: 582-584.

Lyutzkanova, D., J. Distler, J. Altenbuchner (1997) A spectinomycin resistance determinant from the spectinomycin producer *Streptomyces flavopersicus*. Microbiology 143: 2135-2143.

Ma, H., K. Kendall (1994) Cloning and analysis of a gene cluster from *Streptomyces coelicolor* that causes accelerated aerial mycelium formation in *Streptomyces lividans*. J. Bacteriol. 176: 3800-3811.

MacNeil, D. J., K. M. Gewain, C. L. Ruby, G. Dezeny, P. H. Gibbons, T. MacNeil (1992) Analysis of *Streptomyces avermitilis* genes required for avermectin biosynthesis utilizing a novel integration vector. Gene 111: 61-68.

Madduri, K., C. Stuttard, L. C. Vining (1991) Cloning and location of a gene governing lysine ε-aminotransferase, an enzyme initiation β-lactam biosynthesis in *Streptomyces* spp. J. Bacteriol. 173: 985-988.

Malpartida, F., D. A. Hopwood (1984) Molecular cloning of the whole biosynthetic pathway of a *Streptomyces* antibiotic and its expression in a heterologous host. Nature 309: 462-464.

Malpartida, F., J. Niemi, R. Navarrete, D. A. Hopwood (1990) Cloning and expression in a heterologous host of the complete set of genes for biosynthesis of the*Streptomyces coelicolor* antibiotic undecylprodigiosin. Gene 93: 91-99.

Manna, F., D. R. Massardo, K. Wolf, G. Luccarini, M. S. Carlomagno, F. Rivellini, P. Alifano, L. Del Giudice (1994) A tRNA gene mapping within the chloroplast rDNA cluster is differentially expressed during the development of *Daucus carota*. Nucl. Acids Res. 22: 1712-1718.

Mansouri, K., K. Pissowotzki, J. Distler, G. Mayer, P. Heinzel, C. Braun, A. Ebert, W. Piepersberg (1989) Genetics of streptomycin production. *In* Genetics and Molecular Biology of Industrial microorganisms. Edited by C. L. Hershberger, S. W. Queener and G. Hegeman. Washington, D. C.: American Society for Microbiology, pp. 61-67.

Markus, H., B. K. Leskiw (1997) Development of green fluorescent protein-containing promoter probe vectors for use in *Streptomyces*. Biology 499 Project, University of Alberta, Edmonton, Alberta, Canada.

Marsh, E. N., M. D. T. Chang, C. A. Townsend (1992) Two isozymes of clavaminate synthase central to clavulanic acid formation: cloning and sequencing of both genes from from *Streptomyces clavuligerus*. Biochemistry 31: 12648-12657.

Martin, J. F. (1998) New aspects of genes and enzymes for β-lactam antibiotic biosynthesis. Appl. Microbiol. Biotechnol. 50: 1-15.

Martinez-Costa, O. H., P. Arias, N. M. Romero, V. Parro, R. P. Mellado, F. Malpartida (1996) A *relA/spoT* homologous gene from *Streptomyces coelicolor* A3(2) controls antibiotic biosynthetic genes. J. Biol. Chem. 271: 10627-10634.

Matsumoto, A., S.-K. Hong, H. Ishizuka, S. Horinouchi, B. T. (1994) Phosphorylation of the AfsR protein involved in secondary metabolism in *Streptomyces* species by a eukaryotic-type protein kinase. Gene 146: 47-56.

Matsumoto, A., H. Ishizuka, T. Beppu, S. Horinouchi (1995) Involvement of a small ORF downstream of the *afsR* gene in the regulation of secondary metabolism in *Streptomyces coelicolor* A3(2). Actinomycetol. 9: 37-43.

McCormick, J. R., R. Losick (1996) Cell Division gene *ftsQ* is required for efficient sporulation but not growth and viability in *Streptomyces coelicolor* A3(2). J. Bacteriol. 178: 5295-5301.

McCormick, J. R., E. P. Su, A. Driks, R. Losick (1994) Growth and viability of *Streptomyces coelicolor* mutant for the cell division gene *ftsZ*. Mol. Microbiol. 14: 243-254.

McCue, L. A., J. Kwak, M. J. Babcock, K. E. Kendrick (1992) Molecular analysis of sporulation in *Streptomyces griseus*. Gene 115: 173-179.

McCue, L. A., J. Kwak, J. Wang, K. E. Kendrick (1996) Analysis of a gene that suppresses the morphological defect of bald mutants of *Streptomyces griseus*. J. Bacteriol. 178: 2867-2875.

McVittie, A. (1974) Ultrastructural studies on sporulation in wild-type and white colony mutants of *Streptomyces coelicolor*. J. Gen. Microbiol. 81: 291-302.

Merrick, M. J. (1976) A morphological and genetic mapping study of bald colony mutants of *Streptomyces coelicolor*. J. Gen. Microbiol. 96: 299-315.

Messing, J., J. Vieira (1982) A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. Gene 19: 269-276.

Miller, K. (1987) Focus 9: 14-15.

Miyake, K., T. Kuzuyama, S. Horinouchi, T. Beppu (1990) The A-factor-binding protein of *Streptomyces griseus* negatively controls streptomycin production and sporulation. J. Bacteriol. 172: 3003-3008.

Molnar, I., J. F. Aparicio, S. F. Haydock, L. E. Khaw, T. Schwecke, A. Konig, J. Staunton, P. F. Leadlay (1996) Organisation of the biosynthetic gene cluster for rapamycin in *Streptomyces hygroscopicus*: analysis of genes flanking the polyketide synthase. Gene 169: 1-7.

Mori, H., T. Shibasake, K. Yano, A. Ozaki (1997) Purification and cloning of a proline 3-hydroxylase, a novel enzyme which hydroxylates free L-poline to *cis*-3-hydroxy-L-proline. J. Bacteriol. 179: 5677-5683.

Motamedi, H., S.-J. Cai, A. Shafiee, K. O. Elliston (1997) Structural organization of a multifunctional polyketide synthase involved in the biosynthesis of the macrolide immunosuppressant FK506. Eur. J. Biochem. 244: 74-80.

Murgola, E. J. (1995) Translational Suppression: When two wrongs DO make a right. *In:* tRNA: Structure, Biosynthesis, and Function. Edited by D. Soll and U. RajBhandary. Washington, DC: American Society for Microbiology, pp. 491-509.

Natsume, M., K. Yasui, S. Marumo (1989) Calcium ion regulates aerial mycelium formation in Actinomycetes. J. Antibiot. 42: 440-447.

Newman, J. V., R. L. Burghoff, L. Pallesen, K. A. Krogfelt, C. S. Kristensen, D. C. Laux, P. S. Cohen (1994) Stimulation of *Escherichia coli* F-18Col- Type-1 fimbriae synthesis by *leuX*. FEMS Microbiol. Lett. 122: 281-287.

Nihira, T., Y. Shimizu, H. S. Kim, Y. Yamada (1988) Structure-activity relationships of virginiae butanolide C, an inducer of virginiamycin production in *Streptomyces virginiae*. J. Antibiotics 41: 1828-1837.

Nodwell, J. R., R. Losick (1998) Purification of an extracellular signaling molecule involved in production of aerial mycelium by *Streptomyces coelicolor*. J. Bacteriol. 180: 1334-1337.

Nodwell, J. R., K. McGovern, R. Losick (1996) An oligopeptide permease responsible for the import of an extracellular signal governing aerial mycelium formation in *Streptomyces coelicolor*. Mol. Microbiol. 22: 881-893.

Nodwell, J. R., M. Yang, D. Kuo, R. Losick (1999) Extracellular complementation and the identification of additional genes involved in aerial mycelium formation in *Streptomyces coelicolor*. Genetics 151: 569-584.

Ochi, K. (1987a) Changes in nucleotide pools during sporulation of *Streptomyces griseus* in submerged culture. J. Gen. Microbiol. 133: 2787-2795.

Ochi, K. (1987b) Metabolic initiation of differentiation and secondary metabolism by *Streptomyces griseus*: significance of the stringent response (ppGpp) and GTP content in relation to A factor. J. Bacteriol. 169: 3608-3616.

Ochi, K. (1990a) A relaxed *(rel)* mutant of *Streptomyces coelicolor* A3(2) with a missing ribosomal protein lacks the ability to accumultae ppGpp, A-factor and prodigiosin. J. Gen. Microbiol. 136: 2405-2412.

Ochi, K. (1990b) *Streptomyces relC* mutants with an altered ribosomal protein ST-L11 and genetic analysis of a *Streptomyces griseus relC* mutant. J. Bacteriol. 172: 4008-4016.

Ochman, H., M. M. Medhora, D. Garza, D. L. Hartl (1990) Amplification of flanking sequences by inverse PCR. *In* PCR Protocols: A guide to methods and applications. Edited by M. A. Innis, D. H. Gelfand, J. J. Sninsky and T. J. White. San Diego, CA: Academic Press, Inc., pp. 219-227.

Okamoto, S., M. Itoh, K. Ochi (1997) Molecular cloning and characterization of the *obg* gene of *Streptomyces griseus* in relation to the onset of morphological differentiation. J. Bacteriol. 179: 170-179.

Okamoto, S., M. Itoh, K. Ochi (1998) In vitro protein phosphorylation associated with cellular differentiation of *Streptomyces griseus*. Arch Microbiol. 169: 174-177.

Okamoto, S., K. Nakamura, T. Nihira, Y. Yamada (1995) Virginiae butanolide binding protein from *Streptomyces virginiae*. J. Biol. Chem. 270: 12319-12326.

Okamoto, S., K. Ochi (1998) An essential GTP-binding protein functions as a regulator for differentiation in *Streptomyces coelicolor*. Mol. Microbiol. 30: 107-119.

Okamura, K., K. Soga, Y. Shimauchi, T. Ishikura (1977) Holomycin and N-propionylholothin, antibiotics produced by a cephamycin C producer. J. Antibiot. 30: 334-336.

Omer, C. A., S. N. Cohen (1986) Structural analysis of plasmid and chromosomal loci involved in site-specific excision and integration of the SLP1 element of *Streptomyces coelicolor*. J. Bacteriol. 166: 999-1006.

Omer, C. A., R. Lenstra, P. J. Litle, C. Dean, J. M. Tepperman, K. J. Leto, J. A. Romesser, D. P. O'Keefe (1990) Genes for two herbicide-inducible cytochromes P-450 from *Streptomyces griseolus*. J. Bacteriol. 172: 3335-3345.

Onaka, H., N. Ando, T. Nihira, Y. Yamada, T. Beppu, S. Horinouchi (1995) Cloning and characterization of the A-factor receptor gene from *Streptomyces griseus*. J. Bacteriol. 177: 6083-6092.

Onaka, H., S. Horinouchi (1997b) DNA-binding activity of the A-factor receptor protein and its recognition DNA sequences. Mol. Microbiol. 24: 991-1000.

Onaka, H., T. Nakagawa, S. Horinouchi (1998) Involvement of two A-factor receptor homologues in Streptomyces coelicolor A3(2) in the regulation of secondary metabolism and morphogenesis. Mol. Microbiol. 28: 743-753.

Onaka, H., M. Sugiyama, S. Horinouchi (1997a) A mutation at proline-115 in the A-factor receptor protein of *Streptomyces griseus* abolishes DNA-binding ability but not ligand-binding ability. J. Bacteriol. 179: 2748-2752.

Osawa, S., T. H. Jukes (1995) On codon reassignment. J. Mol. Evol. 41: 247-249.

Otten, S. L., J. Ferguson, C. R. Hutchinson (1995) Regulation of daunorubicin production in *Streptomyces peucetius* by the *dnrR*$_2$ locus. J. Bacteriol. 177: 1216-1224.

Paradkar, A. S., K. A. Aidoo, S. E. Jensen (1998) A pathway-specific transcriptional activator regulates late steps of clavulanic acid biosynthesis in *Streptomyces clavuligerus*. Mol. Micorbiol. 27: 831-843.

Paradkar, A. S., K. A. Aidoo, A. Wong, S. E. Jensen (1996) Molecular analysis of a β-lactam resistance gene encoded within the cephamycin gene cluster of *Streptomyces clavuligerus*. J. Bacteriol. 178: 6266-6274.

Paradkar, A. S., S. E. Jensen (1995) Functional analysis of the gene encoding the clavaminate synthase 2 isoenzyme involved in clavulanic acid biosynthesis in *Streptomyces clavuligerus*. J. Bacteriol. 177: 1307-1314.

Passantino, R., A. M. Puglia, K. F. Chater (1991) Additional copies of the *actII* regulatory gene induce actinorhodin production in pleiotropic *bld* mutants of *Streptomyces coelicolor* A3(2). J. Gen. Microbiol. 137: 2059-2064.

Penfold, S. S., J. Simon, L. S. Frost (1996) Regulation of the expression of the *traM* gene of the F sex factor of *Escherichia coli*. Mol. Microbiol. 20: 549-558.

Penyige, A., G. Barabas, I. Szabo, J. C. Ensign (1990) ADP-ribosylation of membrane proteins of *Streptomyces griseus* strain 52-1. FEMS Microbiol. Lett. 69: 293-298.

Penyige, A., G. Vargha, J. C. Ensign, G. Barabas (1992) The possible role of ADP ribosylation in physiological regulation of sporulation in *Streptomyces griseus*. Gene 115: 181-185.

Perez-Llarena, F. J., P. Liras, A. Rodriguez-Garcia, J. F. Martin (1997) A regulatory gene *(ccaR)* required for cephamycin and clavulanic acid production in *Streptomyces clavuligerus*: amplification results in overproduction of both β-lactam compounds. J. Bacteriol. 179: 2053-2059.

Perez-Redondo, R., A. Rodriguez-Garcia, J. F. Martin, P. Liras (1998) The *claR* gene of *Streptomyces clavuligerus* encoding a LysR-type regulatory protein controlling clavulanic acid biosynthesis, is linked to the clavulanate-9-aldehyde reductase *(car)* gene. Gene 211: 311-321.

Perito, B., E. Hanhart, M. Iqbal, A. J. McCarthy, G. Mastromei (1994) Characterization and sequence analysis of a *Streptomyces rochei* A2 endoglucanase-encoding gene. Gene 148: 119-124.

Persson, B. C. (1993) Modification of tRNA as a regulatory device. Mol. Microbiol. 8: 1011-1016.

Peschke, U., H. Schmidt, H.-Z. Zhang, W. Piepersberg (1995) Molecular characterization of the lincomycin-production gene cluster of *Streptomyhces lincolnensis* 78-11. Mol. Microbiol. 16: 1137-1156.

Petrich, A. K., B. K. Leskiw, A. S. Paradkar, S. E. Jensen (1994) Transcriptional mapping of the genes encoding the early enzymes of the cephamycin biosynthetic pathway of *Streptomyces clavuligerus*. Gene 142: 41-48.

Petrich, A. K., X. Wu, K. L. Roy, S. E. Jensen (1992) Transcriptional analysis of the isopenicillin N synthase-encoding gene of *Streptomyces clavuligerus*. Gene 111: 77-84.

Pissowotzki, K., K. Mansouri, W. Piepersberg (1991) Genetics of streptomycin production in *Streptomyces griseus:* molecular structure and putative function of genes *strELMB2N*. Mol. Gen. Genet. 231: 113-123.

Plaskitt, K. A., K. F. Chater (1995) Influences of developmental genes on localized glycogen deposition in colonies of a mycelial prokaryote, *Streptomyces coelicolor* A3(2): a possible interface between metabolism and morphogenesis.
Phil. Trans. R. Soc. Lond. B. 347: 105-121.

Poehling, S., W. Piepersberg, U. F. Wehmeier (1996) Cloning sequencing and regulation of the *secY* gene from *Streptomyces galbus* DSM40480. Unpublished.

Pope, M. K., B. Green, J. Westpheling (1998) The *bldB* gene encodes a small protein required for morphogenesis, antibiotic production, and catabolite control in *Streptomyces coelicolor*. J. Bacteriol. 180: 1556-1562.

Pope, M. K., B. D. Green, J. Westpheling (1996) The *bld* mutants of *Streptomyces coelicolor* are defective in the regulation of carbon utilization, morphogenesis and cell-cell signalling. Mol. Microbiol. 19: 747-756.

Potuckova, L., G. H. Kelemen, K. C. Findlay, M. A. Lonetto, M. J. Buttner, J. Kormanec (1995) A new RNA polymerase sigma factor, σ^F, is required for the late stages of morphological differentiation in *Streptomyces* spp. Mol. Microbiol. 17: 37-48.

Pruess, D. L., M. Kellett (1983) Ro 22-5417, a new clavam antibiotic from *Streptomyces clavuligerus*. I. Discovery and biological activity. J. Antibiot. 36: 208-212.

Puglia, A. M., E. Cappelletti (1984) A *bald* superfertile U.V.-resistant strain in *Streptomyces coelicolor* A3(2). Microbiologica 7: 263-266.

Raibaud, A., A. R. Zalacain, T. G. Holt, R. Tizard, C. J. Thompson (1991) Nucleotide sequence analysis reveals linked *N*-acetyl hydrolase, thioesterase, transport, and regulatory genes encoded by the bialaphos biosynthetic gene cluster of *Streptomyces hygroscopicus*. J. Bacteriol. 173: 4454-4463.

Reading, C., M. Cole (1977) Clavulanic acid: a beta-lactamase-inhibiting beta-lactam from *Streptomyces clavuligerus*. Antimicrob. Agents Chemother. 11: 852-857.

Redenbach, M., H. M. Kieser, D. Denapaite, A. Eichner, J. Cullum, H. Kinashi, D. A. Hopwood (1996) A set of ordered cosmids and a detailed genetic and physical map for the 8 Mb *Streptomyces coelicolor* A3(2) chromosome. Mol. Microbiol. 21: 77-96.

Ritter, A., D. L. Gally, P. B. Olsen, U. Dobrindt, A. Friedrich, P. Klemm, J. Hacker (1997) The Pai-associated *leuX* specific tRNA(Leu,5) affects type 1 fimbriation in pathogenic *Escherichia coli* by control of FimB recombinase expression.
Mol. Microbiol. 25: 871-882.

Robbins, P. W., K. Overbye, C. Albright, B. Benfield, J. Pero (1992) Cloning and high-level expression of chitinase-encoding gene of *Streptomyces plicatus*. Gene 11: 69-76.

Rolinson, G. N. (1982) The history and background of Augmentin.
SA Med. J. 62: 3A-4A.

Romero, J., J. F. Martin, P. Liras, A. L. Demain, N. Rius (1997) Partial purification, characterization and nitrogen regulation of the lysine ε-aminotransferase of *Streptomyces clavuligerus*. J. Ind. Microbiol. Biotechnol. 18: 241-246.

Romero, N. M., V. Parro, F. Malpartida, R. Mellado (1992) Heterologous activation of the actinorhodin biosynthetic pathway in *Streptomyces lividans*.
Nucl. Acids Res. 20: 2767-2772.

Rosenberg, A. H., E. Goldman, J. J. Dunn, F. W. Studier, G. Zubay (1993) Effects of consecutive AGG codons on translation in *Escherichia coli*, demonstrated with a versatile codon test system. J. Bacteriol. 175: 716-722.

Rowley, K. B., R. M. Elford, I. Roberts, W. M. Holmes (1993) *In vivo* regulatory responses of four *Escherichia coli* operons which encode leucyl-tRNAs.
J. Bacteriol. 175: 1309-1315.

Ruan, X., D. Stassi, S. A. Lax, L. Katz (1997) A second type-I PKS gene cluster isolated from *Streptomyces hygroscopicus* ATCC 29253, a rapamycin-producing strain.
Gene 203: 1-9.

Rudd, B. A. M., D. A. Hopwood (1980) A pigmented mycelial antibiotic in *Streptomyces coelicolor*: control by a chromosomal gene cluster. J. Gen. Microbiol. 119: 333-340.

Ryding, N. J., G. H. Kelemen, C. A. Whatling, M. J. Buttner, K. F. Chater (1998) A developmentally regulated gene encoding a repressor-like protein is essential for sporulation in *Streptomyces coelicolor* A3(2). Mol. Microbiol. 29: 343-357.

Sakamoto, K., G. Kawai, T. Niimi, T. Satoh, M. Sekine, Z. Yamaizumi, S. Nishimura, T. Miyazawa, S. Yokoyama (1993) A modified uridine in the first position of the anticodon of a minor species of arginine tRNA, the *argU* gene product, from *Escherichia coli*.
Eur. J. Biochem. 216: 369-375.

Sakata, N., S. Ikeno, M. Hori, M. Hamada, T. Otani (1992) Cloning and nucleotide sequencing of the antitumor antibiotic C-1027 apoprotein gene.
Biosci. Biotech. Biochem. 56: 1592-1595.

Sala-Trepat, J. M., W. C. Evans (1971) The *meta* cleavage of catechol by *Axotobacter* species. Eur. J. Biochem. 20: 400-413.

Sambrook, J., E. F. Fritsch, T. Maniatis (1989) Molecular cloning. A Laboratory Manual. Cold Spring Harbor, NY.: Cold Spring Harbor Laboratory Press.

Sanchez, L., A. F. Brana (1996) Cell density influences antibiotic biosynthesis in *Streptomyces clavuligerus*. Microbiol. 142: 1209-1220.

Sanger, F., S. S. Nicklen, A. R. Coulson (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74: 5463-5467.

Sauer, U., P. Dürre (1992) Possible function of tRNA(Thr,ACG) in regulation of solvent formation in *Clostridium acetobutylicum*. FEMS Microbiol. Lett. 100: 147-154.

Schauer, A. T., A. D. Nelson, J. B. Daniel (1991) Tn*4563* transposition in *Streptomyces coelicolor* and its application to isolation of new morphological mutants. J. Bacteriol. 173: 5060-5067.

Scheu, A.-K., E. Martinez, J. Soliveri, F. Malpartida (1997) *abaB*, a putative regulator for secondary metabolism in *Streptomyces*. FEMS Microbiol. Lett. 147: 29-36.

Schuz, T. C., H. Zahner (1993) Induction of aerial mycelium formation in a *bld* mutant of *Streptomyces tendae* by Borrelidin-a macrolide antibiotic. FEMS Microbiol. Lett. 114: 41-46.

Schwartz, D., R. Alijah, B. Nubbaumer, S. Pelzer, W. Wohlleben (1996) The peptide synthetase gene *phsA* from *Streptomyces viridochromogenes* is not juxtaposed with other genes involved in nonribosomal biosynthesis of peptides. Appl. Environ. Microbiol. 61: 570-577.

Schwecke, T., J. F. Aparicio, I. Holnar, A. Koenig, L. E. Khaw, S. F. Haydock, M. Oliynyk, P. Caffrey, J. Cortes, J. B. Lester, G. A. Boehm, J. Staunton, P. F. Leadlay (1995) The biosynthetic gene cluster for the polyketide immunosuppressant rapamycin. Proc. Natl. Acad. Sci. USA 92: 7839-7843.

Sedlmeier, R., G. Linti, K. Gregor, H. Schmieger (1993) Sequences of tRNA-encoding genes and associated open reading frames of *Streptomyces lividans*. Gene 132: 125-130.

Servin-Gonzalez, L., C. Castro, C. Perez, M. Rubio, F. Valdez (1997) *bldA*-dependent expression of the *Streptomyces exfoliatus* M11 lipase gene (*lipA*) is mediated by the product of a contiguous gene, *lipR*, encoding a putative transcriptional activator. J. Bacteriol. 179: 7816-7826.

Shiina, T., K. Tanaka, H. Takahashi (1991) Sequence of *hrdB*, and essential gene encoding signa-like transcription factor of *Streptomyces coelicolor* A3(2): homology to principal sigma factors. Gene 107: 145-148.

Shima, J., A. Penyige, K. Ochi (1996) Changes in patterns of ADP-ribosylated proteins during differentiation of *Streptomyces coelicolor* A3(2) and its developmental mutants. J. Bacteriol. 178: 3785-3790.

Shpaer, E. (1986) Constraints on codon context in *Escherichia coli* genes: their possible role in modulation the efficiency of translation. J. Mol. Biol. 188: 555-564.

Smardova, J., Z. Dobrova, M. Havranek, J. Janecek (1992) Effect of protein phosphorylation on the activity of RNA polymerase in Streptomycetes. Folia Microbiol. 37: 53-59.

Soliveri, J., E. Vijgenboom, C. Granozzi, K. A. Plaskitt, K. F. Chater (1993) Functional and evolutionary implications of a survey of various actinomycetes for homologues of two *Streptomyces coelicolor* sporulation genes. J. Gen. Microbiol. 139: 2569-2578.

Sorensen, M. A., K. F. Jensen, S. Pedersen (1990) Measurement of translation rates *in vivo* at individual codons and implication of these rate differences for gene expression. *In* Post-Transcriptional Control of Gene Expression, Edited by J. E. G. McCarthy and M. R. Tuite. Springer-Verlag Berlin Heidelberg, NATO ASI *Ser H* 49: 207-216.

Southern, E. M. (1975) Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98: 503-517.

Spanjaard, R. A., K. Chen, J. R. Walker, J. van Duin (1990) Frameshift suppression at tandem AGA and AGG codons by cloned tRNA genes: assigning a codon to *argU* tRNA and T4 $tRNA^{Arg}$. Nucl. Acids Res. 18: 5031-5036.

Spanjaard, R. A., J. van Duin (1988) Translation of the sequence AGG-AGG yields 50% ribosomal frameshift. Proc. Natl. Acad. Sci. USA 85: 7967-7971.

Stapley, E. O., M. Jackson, S. Hernandez, S. B. Zimmerman, S. A. Currie, S. Mochales, J. M. Mata, H. B. Woodruff, D. Hendlin (1972) Cephamycins, a new family of β-lactam antibiotics. I. production by Actinomycetes, including *Streptomyces lactamdurans* sp. Antimicrob. Agents Chemother. 2: 122-131.

Stock, J. B., A. J. Ninfa, A. M. Stock (1989) Protein phosphorylation and regulation of adaptive responses in bacteria. Microbiol. Rev. 53: 450-490.

Sumner-Smith, M., H. Hottinger, I. Willis, T. L. Koch, R. Arentzen, D. Soll (1984) The *sup8* $tRNA^{Leu}$ gene of *Schizosaccharomyces pombe* has an unusual intervening sequence and reduced pairing in the anticodon stem. Mol. Gen. Genet. 197: 447-452.

Susstrunk, U., J. Pidoux, S. Taubert, A. Ullmann, C. J. Thompson (1998) Pleiotropic effects of cAMP on germination, antibiotic biosynthesis and morphological development in *Steptomyces coelicolor*. Mol. Microbiol. 30: 33-46.

Swan, D. G., A. M. Rodriguez, C. Vilches, C. Mendez, J. A. Salas (1994) Characterization of a *Streptomyces antibioticus* gene encoding a type I polyketide synthase which has an unusual coding sequence. Mol. Gen. Genet. 242: 358-362.

Szabo, G., S. Vitalis (1992) Sporulation without aerial mycelium formation on agar medium by *Streptomyces bikiniensis* HH1, an A-factor-deficient mutant. J. Gen. Microbiol. 138: 1887-1892.

Tabor, S., C. C. Richardson (1987) DNA sequence analysis with a modified bacteriophage T7 DNA polymerase. Proc. Natl. Acad. Sci. USA 84: 4767-4771.

Takai, K., N. Horie, Z. Yamaizumi, S. Nishimura, T. Miyazawa, S. Yokayama (1994) Recognition of UUN codons by two leucine tRNA species from *Escherichia coli*. FEBS Lett. 344: 31-34.

Takatsuki, A., G. Tamura (1971) Tunicamycin, a new antibiotic. II. Some biological properties of the antiviral activity of tunicamycin. J. Antibiot. 24: 224-231.

Tanaka, K., H. Aoki, K. Oda, S. Murao, H. Saito, H. Takahashi (1990) Nucleotide sequence of the gene for a metalloproteinase inhibitor of *Streptomyces nigrescens* (SMPI). Nucl. Acids Res. 18: 6433.

Tercero, J. A., J. C. Espinosa, R. A. Lacalle, A. Jimenez (1996) The biosynthetic pathway of the aminonucleoside antibiotic puromycin, as deduced from the molecular analysis of the *pur* cluster of *Streptomyces alboniger*. J. Biol. Chem. 271: 1579-1590.

Thorbjarnardottir, S., T. Dingermann, T. Rafnar, O. S. Andresson, D. Soll, G. Eggertsson (1985) Leucine tRNA family of *Escherichia coli*: nucleotide sequence of the *supP*(Am) suppressor gene. J. Bacteriol. 161: 219-222.

Tillotson, R. D., H. A. B. Wosten, M. Richter, J. M. Willey (1998) A surface active protein involved in aerial hyphae formation in the filamentous fungus *Schizophillum commune* restores the capacity of a bald mutant of the filamentous bacterium *Streptomyces coelicolor* to erect aerial structures. Mol. Microbiol. 30: 595-602.

Tobin, M. B., S. Kovacevic, K. Madduri, J. A. Hoskins, P. L. Skatrud, L. C. Vining, C. Stuttard, J. R. Miller (1991) Localization of the lysine ε-aminotransferase *(lat)* and δ-(L-α-aminoadipyl)-L-cysteinyl-D-valine synthetase *(pcbAB)* genes from *Streptomyces clavuligerus* and production of lysine ε-aminotrasferase activity in *Escherichia coli*. J. Bacteriol. 173: 6223-6229.

Torkkell, S., K. Ylihonko, J. Hakala, M. Skurnik, P. Mantsala (1997) Characterization of *Streptomyces nogalater* genes encoding enzymes involved in glycosylation steps in nogalamycin biosynthesis. Mol. Gen. Genet. 256: 203-209.

Trepanier, N. K., G. D. Nguyen, P. J. Leedell, B. K. Leskiw (1997) Use of polymerase chain reaction to identify a leucyl tRNA in *Streptomyces coelicolor*. Gene 193: 59-63.

Tsao, L.-S., L.-L. Lin, J.-C. Chen, J.-H. Chen, W.-H. Hsu (1993) Cloning and characterization of an α-amylase gene from *Streptomyces lividans*. Biochim. Biophys. Acta 1171: 255-262.

Tsao, S.-W., B. A. M. Rudd, X.-G. He, C.-J. Chang, H. G. Floss (1985) Identification of a red pigment form *Streptomyces coelicolor* A3(2) as a mixture of prodigiosin derivatives. J. Antibiot. 38: 128-131.

Tseng, H.-C., C. W. Chen (1991) A cloned *ompR*-like gene of *Streptomyces lividans* 66 suppresses defective *melC1*, a putative copper-transfer gene. Mol. Microbiol. 5: 1187-1196.

Ueda, K., C.-W. Hsheh, T. Tosake, H. Shinkawa, B. T., S. Horinouchi (1998) Characterization of an A-factor-responsive repressor for *amfR* essential for onset of aerial mycelium formation in *Streptomyces griseus*. J. Bacteriol. 180: 5085-5093.

Ueda, K., K. Miyake, S. Horinouchi, T. Beppu (1993) A gene cluster involved in aerial mycelium formation in *Streptomyces griseus* encodes proteins similar to the response regulators of two-component regulatory systems and membrane translocators. J. Bacteriol. 175: 2006-2016.

Ueda, Y., I. Kumagai, K. Miura (1992) The effects of a unique D-loop structure of a minor tRNA^{Leu} from *Streptomyces* on its structural stability and amino acid accepting activity,. Nucleic Acids Res. 20: 3911-3917.

Umeyama, T., Y. Tanabe, B. D. Aigle, S. Horinouchi (1996) Expression of the *Streptomyces coelicolor* A3(2) *ptpA* gene encoding a phosphotyrosine protein phosphatase leads to overproduction of secondary metabolites in *S. lividans*. FEMS Microbiol. Lett. 144: 177-184.

Urabe, H., H. Ogawara (1995) Cloning, sequencing and expression of serine/threonine kinase-encoding genes from *Streptomyces coelicolor* A3(2). Gene 153: 99-104.

Valentine, B. P., C. R. Bailey, A. Doherty, J. Morris, S. W. Elson, K. H. Baggaley, N. H. Nicholson (1993) Evidence that arginine is a later metabolic intermediate than ornithine in the biosynthesis of clavulanic acid by *Streptomyces clavuligerus*. J. Chem. Soc. Chem. Commun. 1993: 1210-1211.

Vieira, J., J. Messing (1987) Production of single-stranded plasmid DNA. Methods Enzymol. 153: 3-11.

Vogtli, M., P.-C. Chang, S. N. Cohen (1994) *afsR2*: a previously undetected gene encoding a 63-amino-acid protein that stimulates antibiotic production in *Streptomyces lividans*. Mol. Microbiol. 14: 643-653.

Vogtli, M., R. Hutter (1987) Characterisation of the hydroxystreptomycin phosphotransferase gene (*sph*) of *Streptomyces glaucescens*: nucleotide sequence and promoter analysis. Mol. Gen. Genet. 208: 195-203.

Vold, B. S. (1985) Structure and organization of genes for transfer ribonucleic acid in *Bacillus subtilis*. Microbiol. Rev. 49: 71-80.

Vold, B. S., K. Okamoto, B. J. Murphy, C. J. Green (1988a) Transcriptional analysis of *Bacillus subtilis* rRNA-tRNA operons I. The tRNA gene cluster of *rrnB* has an internal promoter. J. Biol. Chem. 263: 14480-14484.

Vomastek, T., R. Nadvornik, J. Janecek, Z. Techinokova, J. Weiser, P. Branny (1998) Characterisation of two putative protein Ser/Thr kinases from actinomycete *Streptomyces granaticolor* both endowed with different properties. Eur. J. Biochem. 257: 55-61.

Vujaklija, D., S. Horinouchi, T. Beppu (1993) Detection of an A-factor-responsive protein that binds to the upstream activating sequence of *strR*, a regulatory gene for streptomycin biosynthesis in *Streptomyces griseus*. J. Bacteriol. 175: 2652-2661.

Vujaklija, D., K. Ueda, S.-K. Hong, T. Beppu, S. Horinouchi (1991) Identification of an A-factor-dependent promoter in the streptomycin biosynthetic gene cluster of *Streptomyces griseus*. Mol. Gen. Genet. 229: 119-128.

Ward, J. M., J. E. Hodgson (1993) The biosynthetic genes for clavulanic acid and cephamycin production occur as a 'super-cluster' in three *Streptomyces*. FEMS Microbiol. Lett. 110: 239-242.

Watanabe, I., F. Nara, N. Serizawa (1995) Cloning, characterization and expression of the gene encoding cytochrome $P\text{-}450_{sca\text{-}2}$ from *Streptomyces carbophilus* involved in production of pravastatin, a specific HMG-CoA reductase inhibitor. Gene 163: 81-85.

Waters, B., D. Vujaklija, M. R. Gold, J. Davies (1994) Protein tyrosine phosphorylation in streptomycetes. FEMS Microbiol. Lett. 120: 187-190.

Weiss, R., D. Lindsley, B. Falahee, J. Gallant (1988) On the mechanism of ribosomal frameshifting at hungry codons. J. Mol. Biol. 203: 403-410.

Weston, L. A., R. J. Kadner (1988) Role of *uhp* genes in expression of the *Escherichia coli* sugar-phosphate transport system. J. Bacteriol. 170: 3375-3383.

White, J., M. Bibb (1997) *bldA* dependence of undecylprodigiosin production in *Streptomyces coelicolor* A3(2) involves a pathway-specific regulatory cascade. J. Bacteriol. 179: 627-633.

Willey, J., R. Santamaría, J. Guijarro, M. Geistlich, R. Losick (1991) Extracellular complementation of a developmental mutation implicates a small sporulation protein in aerial mycelium formation. Cell 65: 641-650.

Willey, J., J. Schwedock, R. Losick (1993) Multiple extracellular signals govern the production of a morphogenetic protein involved in aerial mycelium formation by *Streptomyces coelicolor*. Genes Dev. 7: 895-903.

Williams, J. G., P. J. Mason (1985) Hybridization in the analysis of RNA., Edited by B. D. Hames and S. J. Higgens. Oxford: IRL Press, pp. 139-178.

Wilson, R. K., B. A. Roe (1989) Presence of the hypermodified nucleotide N^6-(Δ^2-isopentenyl)-2-methylthioadenosine prevents codon misreading by *Escherichia coli* phenylalanyl-transfer RNA. Proc. Natl. Acad. Sci. USA 86: 409-413.

Wittmann, S., F. Shareck, D. Kluepfel, R. Morosoli (1994) Purification and characterization of the celB endoglucanase from *Streptomyces lividans* 66 and DNA sequence of the encoding gene. Appl. Environ. Microbiol. 60: 1701-1703.

Wright, F., M. J. Bibb (1992) Codon usage in the G+C-rich *Streptomyces* genome. Gene 113: 55-65.

Wright, H. M., D. A. Hopwood (1976a) Identification of the antibiotic determined by the SCP1 plasmid of *Streptomyces coelicolor* A3(2). J. Gen. Microbiol. 95: 96-106.

Wright, H. M., D. A. Hopwood (1976b) Actinorhodin is a chromosomally-determined antibiotic in *Streptomyces coelicolor* A3(2). J. Gen. Micobiol. 96: 289-297.

Wu, T.-K., R. W. Busby, T. A. Houston, D. B. McIlwaine, L. A. Egan, C. A. Townsend (1995) Identification, cloning, sequencing, and overexpression of the gene encoding proclavaminate amidino hydrolase and characterization of protein function in clavulanic acid biosyntheis. J. Bacteriol. 177: 3714-3720.

Yanisch-Perron, C., J. J. Vieira, J. Messing (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. Gene 33: 103-119.

Ylihonko, K., J. Hakala, T. Kunnari, P. Mantsala (1996a) Production of hybrid anthracycline antibiotic by heterologous expression of *Streptomyces nogalater* nogalamycin biosynthesis genes. Microbiology 142: 1965-1972.

Ylihonko, K., J. Tuikkanen, S. Jussila, L. Cong, P. Mantsala (1996b) A gene cluster involved in nogalamycin biosynthesis from *Streptomyces nogalater:* sequence analysis and complementation of early-block mutations in the anthracycline pathway. Mol. Gen. Genet. 251: 113-120.

Yokoyama, S., S. Nishimura (1995) Modified Nucleosides and Codon Recognition. *In* tRNA: Structure, Biosynthesis, and Function. Edited by D. Soll and U. RajBhandary. Washington, DC: American Society for Microbiology, pp. 207-223.

Zalacain, M., A. Gonzalez, M. C. Guerrero, R. J. Mattaliano, F. Malpartida, A. Jimenez (1986) Nucleotide sequence of the hygromycin B phosphotransferase gene from *Streptomyces hygroscopicus*. Nucl. Acids Res. 14: 1565-1581.

Zhen, L., R. T. Swank (1993) A simple and high yield method for recovering DNA from agarose gels. BioTechniques 14: 894-898.

Zoller, M. J., M. Smith (1987) Oligonucleotide-directed mutagenesis: A simple method using two oligonucleotide primers and a single-stranded DNA template. Methods Enzymol. 154: 329-349.

Zotchev, S. B., H. Schrempf (1994) The linear *Streptomyces* plasmid pBL1: analysis of transfer functions. Mol. Gen. Genet. 242: 374-382.

Zukowski, M. M., D. F. Gaffney, D. Speck, M. Kauffmann, A. Findeli, A. Wisecup, J.-P. Lecocq (1983) Chromogenic identification of genetic regulatory signals in *Bacillus subtilis* based on expression of a cloned *Pseudomonas* gene. Proc. Natl. Acad. Sci. USA 80: 1101-1105.

V1. Appendix 1 : Codon Usage Table of the Genus *Streptomyces*

Codon	Freq*	Codon	Freq*	Codon	Freq*	Codon	Freq*
+UUU Phe F	0.6	+UCU Ser S	0.7	-UAU Tyr Y	1.3	-UGU Cys C	1.0
UUC Phe F	27.4	UCC Ser S	20.3	UAC Tyr Y	19.7	UGC Cys C	7.1
+UUA Leu L	0.2	-UCA Ser S	1.2	UAA OCHRE	0.2	UGA OPAL	2.0
-UUG Leu L	3.6	UCG Ser S	15.2	UAG AMBER	0.4	UGG Trp W	14.1
-CUU Leu L	2.0	-CCU Pro P	1.8	-CAU His H	2.5	CGU Arg R	6.6
CUC Leu L	36.5	CCC Pro P	24.9	CAC His H	22.2	CGC Arg R	36.2
+CUA Leu L	0.5	-CCA Pro P	1.4	-CAA Gln Q	1.9	-CGA Arg R	2.6
CUG Leu L	55.9	CCG Pro P	31.2	CAG Gln Q	26.3	CGG Arg R	29.2
-AUU Ile I	1.0	-ACU Thr T	1.3	-AAU Asn N	1.0	-AGU Ser S	1.8
AUC Ile I	29.3	ACC Thr T	42.5	AAC Asn N	18.8	AGC Ser S	13.4
+AUA Ile I	0.8	-ACA Thr T	2.0	-AAA Lys K	1.1	+AGA Arg R	0.7
AUG Met M	15.7	ACG Thr T	18.7	AAG Lys K	20.5	-AGG Arg R	3.3
-GUU Val V	2.2	-GCU Ala A	4.2	-GAU Asp D	4.8	GGU Gly G	11.1
GUC Val V	45.1	GCC Ala A	76.1	GAC Asp D	58.1	GGC Gly G	58.9
-GUA Val V	2.6	GCA Ala A	5.9	GAA Glu E	10.2	GGA Gly G	7.0
GUG Val V	34.2	GCG Ala A	46.2	GAG Glu E	48.5	GGG Gly G	16.6

Total: 1221 genes

*Frequency per thousand codons

+Rare codons

-Minor codons

Codon usage was tabulated from a compilation of CUTG databases 25/02/1998
Codon usage for individual Streptomyces species can be obtained from the
Online Library at the Streptomyces Internet Resource Center
http://molbio.cbs.umn.edu/asirc/lib/lib.html

VI. APPENDIX II: MODIFIED BASES IN tRNA

mnm	methylaminomethyl
cmnm	carboxymethylaminomethyl
m^{1}A	1-methyladenosine
m^{2}A	methyladenosine
m^{6}A	methylated adenosine
i^{6}A	isopentenyladenosine
t^{6}A	*N*^{6}-threonylcarbamoyladenosine
mt^{6}A	methyl-*N*^{6}-threonylcarbamoyladenosine
I	inosine
m^{1}I	1-methylinosine
Cm	2'-*O*-methylcytidine
ac^{4}C	N^4-acetylcytidine
k^{2}C	lysidine
m^{1}G	1-methylguanosine
m^{7}G	7-methylguanosine
Gm	2'-*O*-methylguanosine
yW	wybutosine
Ψ	pseudouridine
D	dihydrouridine
xm^{5}U	5-methyluridine derivatives (m^{5}U, ribosylthymine; chm^{5}U, 5-(carboxyhydroxymethyl)uridine; mchm^{5}U, 5-(carboxyhydroxymethyl)uridine methyl ester; mcm^{5}U, 5-methoxycarbonylmethyluridine; mnm^{5}U, 5-methylaminomethyluridine; ncm^{5}U, 5-carbamoylmethyluridine; cmnm^{5}U, 5-carboxymethylaminomethyluridine)
Um	2'-*O*-methyluridine
mnm^{5}s^{2}U	5-methylaminomethyl-2-thiouridine
xm^{5}s^{2}U	5-methyl-2-thiouridine derivatives (m^{5}s^{2}U, 5-methyl-2-thiouridine; mcm^{5}s^{2}U, 5-methoxycarbonylmethyl-2-thiouridine; nm^{5}s^{2}U, 5-aminomethyl-2-thiouridine; mnm^{5}s^{2}U, 5-methylaminomethyl-2-thiouridine, cmnm^{5}s^{2}U, 5-carboxymethylaminomethyl-2-thiouridine)
xo^{5}U	5-hydroxyuridine derivatives (ho^{5}U, 5-hydroxyuridine; mo^{5}U, 5-methoxyuridine; cmo^{5}U, uridine 5-oxyacetic acid; mcmo^{5}U, uridine 5-oxyacetic acid methyl ester)
xm^{5}Um	5-methyl-2'-*O*-methyluridine derivatives (m^{5}Um, 5,2'-*O*-dimethyluridine; mcm^{5}Um, 5-methoxycarbonylmethyl-2'-*O*-methyluridine, ncm^{5}Um, 5-carbamoylmethyl-2'-*O*-methyluridine, cmnm^{5}Um, 5-carboxymethylaminomethyl-2'-*O*-methyluridine)